

ON SOME GENERALIZATIONS OF SKEW-SHIFTS ON \mathbb{T}^2

KRISTIAN BJERKLÖV

ABSTRACT. In this paper we investigate maps of the two-torus \mathbb{T}^2 of the form $T(x, y) = (x + \omega, g(x) + f(y))$ for Diophantine $\omega \in \mathbb{T}$ and for a class of maps $f, g : \mathbb{T} \rightarrow \mathbb{T}$, where each g is strictly monotone and of degree 2, and each f is an orientation preserving circle homeomorphism. For our class of f and g we show that T is minimal and has exactly two invariant and ergodic Borel probability measures. Moreover, these measures are supported on two T -invariant graphs. One of the graphs is a Strange Nonchaotic Attractor whose basin of attraction consists of (Lebesgue) almost all points in \mathbb{T}^2 . Only a low regularity assumption (Lipschitz) is needed on the maps f and g , and the results are robust with respect to Lipschitz-small perturbations of f and g .

MSC2010: 37C40, 37C70, 37E30

1. INTRODUCTION

Skew-shifts are classical examples in ergodic theory¹. These are maps of the two-torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ of the form

$$T_0(x, y) = (x + \omega, kx + y)$$

where $k \in \mathbb{Z} \setminus \{0\}$ and $\omega \in \mathbb{R} \setminus \mathbb{Q}$. It is well-known that they are minimal and uniquely ergodic (see, e.g., [9, Prop. 4.7.4]). A natural question is to ask what can happen with the dynamics of T_0 if one changes the fiber map $y \mapsto kx + y$ to, e.g., $y \mapsto kx + f(y)$, where f is a circle homeomorphism; this is the topic of the present paper. In particular we will be interested in the case $k = 2$ which is the simplest case where one gets "interacting critical sets", or "resonances"². This case also naturally arises in certain models (see Example 2 below). Showing how such cases can be handled for certain classes of f , without using any parameter exclusion, and under low regularity assumptions, is one of the main purposes of this paper. The case $k = 1$ is easier (but far from understood for general f) and becomes a special case of our analysis (for the class of maps f under consideration; see Subsection 1.2). Other existing techniques (for example, [5, 28]) could also have been adapted for analyzing the case $k = 1$, since there is no need of any exclusion of parameters in this situation (as already noticed in [28, Remark 2]).

Research supported by the Swedish Research Council, 2012-3090.

¹They are also used, for example, as base dynamics in the theory of ergodic Schrödinger operators. See, e.g., [7].

²This is due to the fact every circle $\mathbb{T} \times \{y\}$, $y \in \mathbb{T}$, is mapped to a closed curve which intersects each circle $\mathbb{T} \times \{\eta\}$ twice.

Skew-shifts are examples of so-called quasi-periodically driven circle maps (skew-product maps), i.e., maps on \mathbb{T}^2 of the form

$$F(x, y) = (x + \omega, h(x, y))$$

where each map $\mathbb{T} \ni y \mapsto h_x(y) = h(x, y) \in \mathbb{T}$ is a circle homeomorphism. This (large) class of maps (which all have zero topological entropy; see [8]) have been intensely investigated in the literature (see, e.g., [2, 4, 5, 14, 18, 20, 25, 28], and references therein), under various assumption on $h(x, y)$; but still a general theory seems very remote³. One of the central problems that arises when analyzing such maps is the presence of "resonances". They appear in one way or another (and on infinitely many scales), for examples in the form of so-called small divisors. There are essentially two ways to handle this problem: either to face it, or try to avoid it. A way to avoid resonances is to consider parameter families of maps F_a , and use some kind of parameter exclusion to end up with a "large" set of parameter values a for which one can describe the dynamics of F_a . This method has been used, e.g., in [11]. It is also the route taken in, for example, [4, 18, 19, 28] (based on ideas from one-dimensional dynamics, e.g., [3, 10, 17]). The price one typically has to pay to be able to treat a fixed map F (from some class of maps), or to be able to say something about the dynamics of F_a for *all* parameter values a , is to put more assumptions on the map $h(x, y)$ (and often work harder). Examples of the latter case are [5, 14, 24, 25]. In this paper we shall put a topological condition on $h(x, y)$ (as, e.g., in [2, 27]).

In the case of skew-shifts we have $h(x, y) = kx + y$. Note that the map $g(x) = kx$ has degree $k \neq 0$, and thus T_0 is not homotopic to the identity map. A big difference between maps F that are homotopic to the identity, and those which are not, is that the latter ones cannot have continuous invariant curves, i.e., there are no continuous functions $\varphi : \mathbb{T} \rightarrow \mathbb{T}$ such that $F(x, \varphi(x)) = (x + \omega, \varphi(x + \omega))$ for all $x \in \mathbb{T}$ (see, e.g., [14, Prop.4.2]). However, there can be (highly discontinuous) *measurable* functions $\varphi : \mathbb{T} \rightarrow \mathbb{T}$ whose graphs are almost everywhere F -invariant. If such a "strange" invariant graph (which can, and do, also arise in maps F homotopic to the identity) attracts points, it is often called a Strange Nonchaotic Attractor (SNA). Such objects has shown to exist in various quasi-periodically driven models (see, e.g., [4, 5, 14, 16, 18, 28], and references therein). There is also a huge literature focusing on SNAs based on numerical experiments.

A way to measure contraction (note that there can only be contraction in the y -direction, since we have rotation in the x -direction) is via the (fibered) Lyapunov exponents. Given a point $(x, y) \in \mathbb{T}^2$ we use the notation

$$(x_n, y_n) = F^n(x, y), \quad n \in \mathbb{Z}.$$

³The situation where the base dynamics $x \mapsto x + \omega$ is an ergodic translation on \mathbb{T}^d , $d \geq 2$, still lacks of techniques for rigorous analysis in many interesting cases, including our setting. Furthermore, the case of continuous maps on \mathbb{T}^2 , or $\mathbb{T} \times I$ (I an interval), of the form $F(x, y) = (x + \omega, q(x, y))$ where the maps $q(x, \cdot)$ are not necessarily invertible is much less understood; see, e.g., [6] for some results in this direction.

Assuming, for example, that $h : \mathbb{T}^2 \rightarrow \mathbb{T}$ is C^1 , we can define the (fibered) Lyapunov exponents

$$\lambda(x, y) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log \left| \frac{\partial y_n}{\partial y} \right| = \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \log \left| \frac{\partial h}{\partial y}(x_k, y_k) \right|.$$

For the map T_0 we trivially have $\lambda(x, y) = 0$ for all $(x, y) \in \mathbb{T}^2$. Furthermore, since the map F is invertible, we can have $\lambda(x, y) > 0$ for, at most, a set of zero Lebesgue measure of $(x, y) \in \mathbb{T}^2$. Thus, the question is whether $\lambda(x, y) = 0$ for a.e. $(x, y) \in \mathbb{T}^2$, or if $\lambda(x, y) < 0$ for a positive measure set of $(x, y) \in \mathbb{T}^2$. Unfortunately this question is often notoriously hard to answer, but central for the description of the dynamics.

1.0.1. Perturbations of skew-shifts. We return to the skew-shift T_0 . It has been noticed that the dynamics of this map seems to be sensitive to perturbations (in the fiber), but that the perturbed maps have a more robust dynamics. More precisely, in [20] the authors investigate maps of \mathbb{T}^2 of the form

$$F(x, y) = (x + \omega, x + y + \eta P(x, y))$$

where $P : \mathbb{T}^2 \rightarrow \mathbb{R}$ is C^1 , and the parameter η ranges in an interval for which F is invertible. An example considered in the numerical experiments in [20] is $P(x, y) = \sin(2\pi y)$. These experiments indicate that the map F has an attractor with a negative (fibered) Lyapunov exponent for all (small) $\eta \neq 0$. Whether this is true or not seems to be an open question. However, under such an assumption, it is rigorously shown in [20] that the map F has a single attractor, and F is topologically transitive. Thus, assuming that the Lyapunov exponent is negative for an open set of parameters η one gets robust strange nonchaotic attractors, meaning that the SNA survives after perturbation (in many other models it is expected that the SNA is very sensitive to perturbations, existing only on the boundary between order and chaos). See Figure 1 for a plot of an orbit of a similar kind of map.

In this paper we will show that such a situation, with robust SNAs, indeed can occur in models included in the wide class considered in [20]. See Theorem 1 below.

In fact, robust families of SNAs also follow by considering the projective action of the non-uniformly hyperbolic $SL(2, \mathbb{R})$ -cocycles (see next paragraph for more details on quasi-periodic cocycles) in [28, Corollary 1], combined with [14, Proposition 4.17]. See also Example 2 below for an extension of Young's result [28, Corollary 1].

1.0.2. Cocycles. A special class of maps F , and a well-studied one, arises in quasi-periodic $SL(2, \mathbb{R})$ cocycles. Consider maps $\Phi : \mathbb{T} \times \mathbb{R}^2 \rightarrow \mathbb{T} \times \mathbb{R}^2$

$$\Phi(x, v) = (x + \omega, M(x)v)$$

where $M : \mathbb{T} \rightarrow SL(2, \mathbb{R})$ is continuous. The cocycle Φ induces a projective action, denoted \overline{F} , on $\mathbb{T} \times \mathbb{P}^1$. If $v \in \mathbb{R}^2 \setminus \{0\}$, we denote by $\overline{v} \in \mathbb{P}^1$ the projectivization of v . We introduce coordinates on $\mathbb{P}^1 = \mathbb{T}$ by letting $\overline{\begin{pmatrix} \cos \pi y \\ \sin \pi y \end{pmatrix}}$ correspond to $y \in \mathbb{T}$. Moreover, an invertible

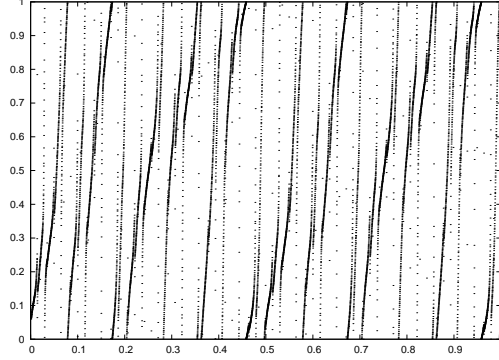


FIGURE 1. Plot of (a piece) of the orbit of $(0,0)$ under the map $F(x,y) = \left(x + \frac{1}{\sqrt{2}}, 2x + y + 0.1 \sin 2\pi y\right)$.

matrix M induces a map \overline{M} on \mathbb{P}^1 in the natural way. Seen like this, the cocycle Φ induces a quasi-periodically driven circle map \overline{F} . For example, if $M(x) = R_{2\pi x}$, where

$$(1.1) \quad R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix},$$

then it is easy to check that $\overline{F}(x,y) = (x + \omega, 2x + y)$, i.e., we get T_0 with $k = 2$.

To the cocycle Φ there is associated the (maximal) Lyapunov exponent

$$L(M) = \lim_{n \rightarrow \infty} \frac{1}{n} \int_{\mathbb{T}} \log \|M^n(x)\| dx \geq 0$$

where $M^n(x) = M(x + (n-1)\omega) \cdots M(x)$, $n \geq 1$. From Oseledets' theorem, it follows that if $L(M) > 0$, then there exists a measurable splitting $W^+(x) \oplus W^-(x) = \mathbb{R}^2$ such that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log |M^n(x)v| = L(M) \text{ for a.e. } x \in \mathbb{T}, \text{ all } v \notin W^-(x).$$

We note that if we let $v_n = M^n(x)v$ and $w_n = M^n(x)w$, where $v, w \notin W^-(x)$ and $\bar{v} \neq \bar{w}$, then, since $M^n(x) \in SL(2, \mathbb{R})$, we have (by considering area)

$$|v_n||w_n| \sin \theta_n = |v||w| \sin \theta$$

where $\theta_n \in [0, \pi)$ is the angle between v_n and w_n . For a.e. $x \in \mathbb{T}$ we then have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \sin \theta_n = -2L(M).$$

Thus, we obtain the Lyapunov exponent $L(M)$ in the contraction rate (in the fiber) in iterates of \overline{F} .

The by far most studied case is the one-parameter family

$$M(x) = M_E(x) = \begin{pmatrix} 0 & 1 \\ -1 & q(x) - E \end{pmatrix}, \quad E \in \mathbb{R},$$

where $q : \mathbb{T} \rightarrow \mathbb{R}$. This family of cocycles arises when investigating the one-frequency quasi-periodic Schrödinger equation $-(u_{n+1} + u_{n-1}) + q(x + n\omega)u_n = Eu_n$. In this case the cocycles are homotopic to the identity. For real-analytic q we are beginning to get a good understanding of many properties of the cocycles (see, for example, [1] and references therein). On the contrary, if q is only assumed to have finite smoothness, then the results are much more sparse. If one uses parameter exclusion (in E) there are results in, e.g., [5]. If we let $q(x) = Kq_0(x)$ where $q_0 : \mathbb{T} \rightarrow \mathbb{R}$ is C^2 and has exactly two non-degenerate critical points ("cosine-like"), then it is possible to describe much of the dynamics of Φ_E (or \overline{F}_E) for all values of E , provided that the constant K is large and ω is Diophantine [4, 25] (spectral properties of the associated Schrödinger operator, for this setting, is treated in [12, 23]). For example, one only gets SNA (for the associated projective map \overline{F}) for parameters E in a Cantor set (of large measure) [22, 23, 26]; thus they are not robust. However, for a fixed q_0 of finite smoothness with more than two monotonicity intervals it is not known what can happen; one would have to improve the techniques in the proofs in [4, 25], since one in this case can have "multiple resonances" (although this is not "typical").

Recently there has been an increased interest in quasi-periodic cocycles non-homotopic to the identity. In [2] it is shown that the (smooth) non-homotopic case (or, more precisely, the so-called monotonic case) is surprisingly well-behaved and much more robust than, for example, the Schrödinger case. The results in the present paper is also an indication of this (the examples in [27] are also related to this), that "monotonic" maps are more robust (for example, we can work with low regularity assumptions, even though we have resonances). In [2] it is also shown that the projective action of so-called pre-monotonic cocycles is always minimal. Only $C^{1+\varepsilon}$ smoothness is needed for the argument (which is partly based on an argument from [20]). An interesting problem is whether this holds under even lower regularity assumptions (see also [2, Problem 3.1]).

1.1. Main object. In this paper we shall consider maps $T : \mathbb{T}^2 \rightarrow \mathbb{T}^2$ of the form⁴

$$(1.2) \quad T(x, y) = (x + \omega, g(x) + f(y))$$

where the functions $f, g : \mathbb{T} \rightarrow \mathbb{T}$ are assumed to be Lipschitz continuous.

We recall that a function $\varphi : \mathbb{T} \rightarrow \mathbb{T}$ is Lipschitz continuous if there exists a constant K such that $d(\varphi(y), \varphi(x)) \leq Kd(x, y)$ for all $x, y \in \mathbb{T}$; here d denotes the standard metric on \mathbb{T} . The constant K is called a Lipschitz constant for φ . We say that φ has a Lipschitz constant c on $I \subset \mathbb{T}$ if $d(\varphi(y), \varphi(x)) \leq cd(x, y)$ for all $x, y \in I$. Recall that a Lipschitz continuous function is absolutely continuous, and thus differentiable a.e. on \mathbb{T} .

We also recall that if $\varphi : \mathbb{T} \rightarrow \mathbb{T}$ is a continuous function, then a lift of φ is a continuous function $\Phi : \mathbb{R} \rightarrow \mathbb{R}$ such that $\pi \circ \Phi(x) = \varphi \circ \pi(x)$ for all $x \in \mathbb{R}$, where $\pi : \mathbb{R} \rightarrow \mathbb{T}$ is the natural projection $\pi(x) = x \bmod 1$. It is well-known that a lift of φ is unique up to the addition of an integer. Moreover, $\Phi(x+1) - \Phi(x) = k$, $k \in \mathbb{Z}$. The integer k is called the degree of φ .

⁴We could have considered more general classes of maps, i.e., that we have the exact form $g(x) + f(y)$ is not important for the method we use; but for simplicity we have chosen to work with this class.

Note that for maps T of the form (1.2), the (fibered) Lyapunov exponents $\lambda(x, y)$ become

$$\lambda(x, y) = \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \log |f'(y_k)|.$$

1.2. Assumptions on f, g and ω . Here we specify the exact assumptions on f and g . We begin with the assumption on g :

(A1)(κ) *We assume that $g : \mathbb{T} \rightarrow \mathbb{T}$ has degree 2, and that a lift $G : \mathbb{R} \rightarrow \mathbb{R}$ of g is bi-Lipschitz, and G and G^{-1} both have a Lipschitz constant $< \kappa$.*

Remark 1. An natural example is $g(x) = 2x$. Our result (Theorem 1) also holds in the easier case when g has degree 1, for example if $g(x) = x$. The proof is easier (we do not get any "resonant cases"; see Section 2.3, and Sections 5 and 6). However, one of the main purposes of this paper is to treat "resonances", so we do not stress on this case.

In our analysis it would in fact be possible to allow g to have any degree $k \geq 1$ (and not adding any smoothness assumptions). However, the bookkeeping (and notation) in the proof of Theorem 1 below would become slightly more involved; but fundamentally no new problems would arise. Thus we have chosen to concentrate on the case $k = 2$.

In the assumptions on f below, the number $\varepsilon > 0$ should be thought of being a (very) small number.

(A2)(ε, ρ) *Assume that $f : \mathbb{T} \rightarrow \mathbb{T}$ is an orientation preserving bi-Lipschitz homeomorphism, where f and f^{-1} both have a Lipschitz constant $< \varepsilon^{-\rho}$. We further assume that there exist disjoint open intervals $A, B \subset \mathbb{T}$ such that*

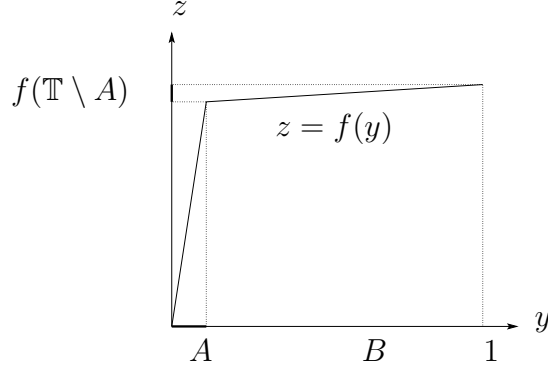
$$|\mathbb{T} \setminus B| < \varepsilon, |f(\mathbb{T} \setminus A)| < \varepsilon$$

and f has a Lipschitz constant $< \varepsilon$ on B , and f^{-1} has a Lipschitz constant $< \varepsilon$ on $f(A)$.

That f is bi-Lipschitz implies that f and f^{-1} both are differentiable for a.e. $y \in \mathbb{T}$. We note that the two last conditions imply

$$f'(y) < \varepsilon \text{ for a.e } y \in B; \quad f'(y) > \varepsilon^{-1} \text{ for a.e } y \in A.$$

Remark 2. We have tried to make the assumptions on f as clean as possible, avoiding the introduction of many parameters. It would be possible to have other assumptions. However, we do really need that the assumptions are "perturbative", in the sense that the ε is very small. This will enable us, in the proof, to start the induction immediately. Without such a smallness assumption one would need to develop some finer analysis on the base case, which should give estimates (similar to the ones we have put on f) for some iterate T^k . Such an approach would be very interesting. However, here we focus on one part of the main mechanisms.

FIGURE 2. The function f in example 1.

Example 1. An example of a map satisfying $(A2)(\varepsilon, 2)$ is

$$f(y) = \begin{cases} 2y/\varepsilon, & 0 \leq y \leq a \\ \varepsilon(y-1)/2 + 1, & a \leq y \leq 1 \end{cases}$$

where $a = (2\varepsilon - \varepsilon^2)/(4 - \varepsilon^2)$. Here we can take $A = (0, a)$ and $B = (a, 1)$. See Figure 2

Remark 3. Note that the assumptions on f and g are robust in the following sense: If g satisfy $(A1)(\kappa)$ and f satisfy $(A2)(\varepsilon, \rho)$ (for some $\varepsilon, \kappa, \rho > 0$), and if $f_1, g_1 : \mathbb{T} \rightarrow \mathbb{T}$ are Lipschitz continuous with a Lipschitz constant sufficiently small, then $g + g_1$ satisfy $(A1)(\kappa)$ and $f + f_1$ satisfy $(A2)(\varepsilon, \rho)$. Note also that if g satisfy $(A1)(\kappa)$, then, trivially, $g(x) + t$ also satisfy $(A1)(\kappa)$ for any $t \in \mathbb{T}$.

Next follows our assumption on the base frequency ω . We assume that ω satisfies the Diophantine condition ⁵

$$(1.3) \quad \inf_{p \in \mathbb{Z}} |q\omega - p| > \frac{\gamma}{q^2} \quad \text{for all } q \in \mathbb{Z} \setminus \{0\},$$

for some constant $\gamma > 0$. Notice that this condition is satisfied for (Lebesgue) a.e. ω .

1.3. Statement of results. We are now ready to state our results for maps T defined as in (1.2):

Theorem 1. *Assume that ω satisfies the Diophantine condition (1.3) for some $\gamma > 0$. Given any numbers $\kappa, \rho > 1$ there is an $\varepsilon_0 = \varepsilon_0(\gamma, \kappa, \rho) > 0$ such that for all $0 < \varepsilon \leq \varepsilon_0$ the following holds for the map $T(x, y) = (x + \omega, g(x) + f(y))$ whenever g satisfies $(A1)(\kappa)$ and f satisfies $(A2)(\varepsilon, \rho)$:*

- (1) *The map T is minimal;*
- (2) *The Lyapunov exponent $\lambda(x, y) < (\log \varepsilon)/2$ for a.e. $(x, y) \in \mathbb{T}^2$;*

⁵We could equally well have taken the lower bound of the form $\gamma/|q|^{1+\tau}$, any $\tau \geq 1$; but in order not to have many constants we have chosen to work with q^2 . In fact one could even have a weaker condition, like Brjuno, as, for example, in [28].

- (3) T has exactly two invariant and ergodic Borel probability measures μ^s, μ^u . Moreover, μ^s, μ^u are the push-forward of the Lebesgue measure on \mathbb{T} by two maps $x \mapsto (x, u(x))$ and $x \mapsto (x, s(x))$, respectively, where $u, s : \mathbb{T} \rightarrow \mathbb{T}$ are distinct measurable functions whose graphs are a.e. T -invariant.
- (4) For a.e. $x \in \mathbb{T}$ and all $y \neq s(x)$ we have $d(y_n, u(x_n)) < C(x, y)\varepsilon^{n/2}$, $n \geq 0$.

Remark 4. a) Condition (4) tells us that most points (x, y) are attracted, under forward iteration by T , to the graph of the measurable function $\xi \mapsto u(\xi)$.

b) We note that if f is also C^1 , then, since we by Birkhoff's ergodic theorem have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \log |f'(y_k)| = \int_{\mathbb{T}^2} \log |f'| d\mu^u \text{ for } \mu^u\text{-a.e. } (x, y) \in \mathbb{T}^2,$$

it follows from (3) and (4) that

$$\lambda(x, y) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \log |f'(y_k)| = \int_{\mathbb{T}} \log |f'(u(x))| dx \text{ for a.e. } (x, y) \in \mathbb{T}^2.$$

Example 2. We can apply the above theorem to a class of $SL(2, \mathbb{R})$ -cocycles over Diophantine rotation. Consider the cocycle

$$\Phi(x, v) = (x + \omega, R_{2\pi\varphi(x)} Dv)$$

where ω satisfies (1.3) for some $\gamma > 0$, D is the diagonal matrix

$$D = \begin{pmatrix} K & 0 \\ 0 & 1/K \end{pmatrix}, \quad K > 1,$$

R_θ is the rotation matrix defined in (1.1), and $\varphi : \mathbb{T} \rightarrow \mathbb{T}$ is any bi-Lipschitz map. Let $L(K)$ denote the (maximal) Lyapunov exponent of the cocycle (recall the discussion in Subsection 1.0.2 above).

It is easy to check that the projectivization of Φ is

$$\overline{F}(x, y) = (x + \omega, 2\varphi(x) + f(y))$$

where $f(y) = \arctan(\tan(\pi y)/K^2)/\pi \pmod{1}$. Note that $g(x) := 2\varphi(x)$ has degree 2, and let $\kappa > 1$ be a Lipschitz constant for G and G^{-1} , where G is a lift of g . Thus g satisfy condition (A1)(κ). Furthermore, given a small $\varepsilon > 0$ it is easy to verify that $f(y)$ satisfies (A2)($\varepsilon, 2$) (i.e., $\rho = 2$) provided that the constant K is sufficiently large. Hence, we can apply Theorem 1 (with the given γ, κ and $\rho = 2$) and conclude that the projective action \overline{F} is minimal and that the (fibered) Lyapunov exponent $\lambda(x, y) < \text{const.} \log(1/K)$, and thus $L(K) > \text{const.} \log K$, for all sufficiently large K (depending on γ and κ).

1.4. Outline of the paper. The rest of the paper is organized as follows: In Section 2 we introduce notations and define some important sets related to the map T . We also derive basic properties of iterations of T . In the last part of Section 2 we derive a few elementary topological properties which we shall use to control the geometry and topology of iterations of curves. Thus, they are central for our analysis. In Section 3 we begin by briefly discussing properties of Lipschitz continuous functions. We also derive formulae for

the derivative of the iterates, and use them to obtain quantitative estimates under suitable assumptions on the iterates. Theorem 1 is proved in Section 4, assuming that the main proposition, Proposition 4.1, holds. The latter proposition gives a very detailed description of the dynamics of T (under the given assumptions). In the following two sections, Section 5 and 6, Proposition 4.1 is proved via an inductive scheme. These two sections contain the key part of the analysis. In Section 5 we verify the base case of the construction. The general approach for analyzing the map T is similar to the ones we have used in [4, 5] (see also [18, 19] where the methods in [5] are adapted to larger classes of circle maps). But in the present paper we exploit the fact that we have maps with different topological properties (as in, for example, [2]).

The (inductive) analysis consists of two main parts: an arithmetic one, and a geometric/topologic one. The arithmetic part is the more standard one, and is very similar to the one we used in [4]. The structure of several of the estimates that arise in this part are rather general and naturally arise in the analysis of non-uniformly hyperbolic maps. However, in the present case, compared with, for example, [4, 5], the topology is very different. This allows us to handle the "resonances" without using parameter exclusion (as is needed in [5, 18, 28]) and without "losing derivatives" (which one does, for example, in [4]). In the non-resonant case, similar techniques were developed by Young in [28]. All these techniques, in turn, are based on ideas from one-dimensional dynamics, in particular [3, 10].

In the final section, Section 8, we have collected "abstract" results, most of them of a computational nature, which are used in different parts of our construction.

2. PRELIMINARIES

In this section we introduce some of the notations which we will be using. We also derive elementary properties of the map T (defined in (1.2)), assuming that g and f satisfy (A1) and (A2), respectively. In the last subsection we will derive some elementary topological facts which will be frequently used to control the geometry.

From now on we assume that ω is fixed, satisfying (1.3) for some $\gamma > 0$. Furthermore, we assume that $\kappa, \rho > 1$ are fixed, and we fix g satisfying condition (A1)(κ). We also assume that f satisfy condition (A2)(ε, ρ), where $0 < \varepsilon \ll 1$, and the smallness of ε is only allowed to depend on κ, ρ and γ .

2.1. Some notations.

- By d we denote the standard metric on \mathbb{T} .
- If $X, Y \subset \mathbb{T}$ are sets, we define the distance $\text{dist}(X, Y) = \inf_{x \in X, y \in Y} d(x, y)$.
- If $I \subset \mathbb{T}$, we denote the interior of I by $\text{int}(I)$; by $|I|$ we denote the (Haar) measure of the set I (provided that I is measurable).
- If $I = [x - \delta, x + \delta] \subset \mathbb{T}$ is an interval, we denote by $3I$ the interval $3I = [x - 3\delta, x + 3\delta]$. Thus, the two intervals have the same center, and $|3I| = 3|I|$ (if I is sufficiently small).
- We let $\pi : \mathbb{R} \rightarrow \mathbb{T}$ be the natural projection $\pi(x) = x \bmod 1$.

- If $y, \eta \in \mathbb{T}$, the (positively oriented) interval (or arc) $[y, \eta]$ is defined in the natural way: Let $\tilde{y}, \tilde{\eta} \in [0, 1)$ be the unique points such that $\pi(\tilde{y}) = y$ and $\pi(\tilde{\eta}) = \eta$. If $\tilde{y} \leq \tilde{\eta}$, then $[y, \eta] = \pi([\tilde{y}, \tilde{\eta}])$; if $\tilde{y} > \tilde{\eta}$, then $[y, \eta] = \pi([\tilde{y}, 1 + \tilde{\eta}])$.
- If $(x, y) \in \mathbb{T}^2$, then we define the projections π_1, π_2 by

$$\pi_1(x, y) = x, \quad \pi_2(x, y) = y.$$

- We say that a function $f : I \rightarrow \mathbb{T}$ ($I \subset \mathbb{T}$ an interval) is (strictly) increasing if the lifts of f are (strictly) increasing.
- If $I \subset \mathbb{T}$ is any set and $x \in \mathbb{T}$, we let $N(x; I)$ denote the smallest integer $k \geq 0$ such that $x + k\omega \in I$. If no such integer exists, we let $N(x; I) = \infty$. Note that $N(x; I) = 0$ if $x \in I$. Thus, N is the first entry time of x to I .
- If $r > 0$ is a real number, we denote by $[r]$ the integer part of r . Thus $r - 1 < [r] \leq r$.

2.2. Definition of the sets A', A'', R and the point β . We now define some sets which play a central rôle in the analysis. Let

$$R = f(\mathbb{T} \setminus A).$$

Since A is open, we have that R is a closed interval. Moreover, by assumption (A2) we have $|R| < \varepsilon$.

By assumption (A2) we have $A \subset \mathbb{T} \setminus B$, and thus $|A| < \varepsilon$. To have some space around the interval A , we let A' be the closed set

$$(2.1) \quad A' = \{y \in \mathbb{T} : \text{dist}(y, A) \leq \varepsilon\}.$$

Since $|A| < \varepsilon$ we have $|A'| < 3\varepsilon$. Moreover, since we clearly have $|A'| > 2\varepsilon$, and since $A \subset A'$, it follows that

$$(2.2) \quad \overline{\mathbb{T} \setminus A'} \subset B.$$

At a few occasions we will also need an intermediate set, A'' , which we define by

$$(2.3) \quad A'' = \{y \in \mathbb{T} : \text{dist}(y, A) \leq \varepsilon/2\}.$$

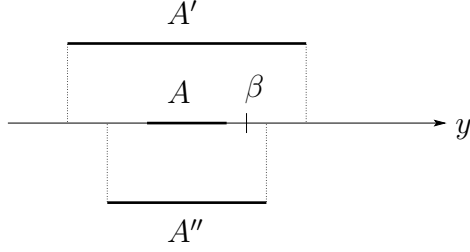
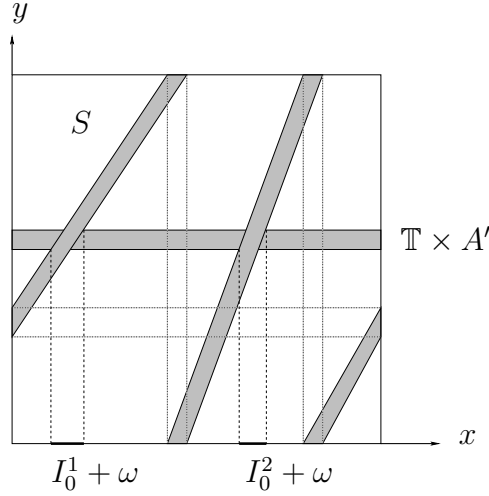
By definition we thus have $A \subset A'' \subset A'$.

We will also use a "reference point" $\beta \in \mathbb{T} \setminus A$ close to A . Writing $A = (a_1, a_2)$, we let

$$(2.4) \quad \beta = a_2 + \varepsilon/4 \in A''.$$

From this choice we immediately get the following (see Figure 3, and recall the definition of the oriented intervals $[y, \eta]$ above):

Lemma 2.1. *If $y \in \mathbb{T} \setminus A''$ then $[\beta, y] \subset \mathbb{T} \setminus A$.*

FIGURE 3. The sets A , A' and A'' , and the point β FIGURE 4. The strips S and $\mathbb{T} \times A'$, and the two intervals $I_0^1 + \omega$ and $I_0^2 + \omega$.

2.3. Definition of the strip S and the set I_0 . We recall that the map T is given by

$$T(x, y) = (x + \omega, g(x) + f(y)).$$

By using the definition of the (small) interval R above, we see that if $(x, y) \in \mathbb{T} \times (\mathbb{T} \setminus A)$, then $T(x, y) \in \{x + \omega\} \times (R + g(x))$. Thus, as long as the iterates $T^k(x, y) \in \mathbb{T} \times (\mathbb{T} \setminus A)$ (for $k \geq 0$), we know that they must be in the strip S defined as

$$(2.5) \quad S = \bigcup_{x \in \mathbb{T}} \{x\} \times (R + g(x - \omega)) = \bigcup_{x \in \mathbb{T}} \{x + \omega\} \times (R + g(x)).$$

If the iterates stayed in this strip for ever (which they cannot), we would easily have a good control on the dynamics. However, sooner or later the iterates will enter the strip $\mathbb{T} \times A$; and on A the map f is strongly expanding (and $|f(A)|$ is close to 1), and thus uncertainty is added. Indeed, since $g(\mathbb{T}) = \mathbb{T}$ we see that there must be a connection between the strip S and the strip $\mathbb{T} \times A$ (see Figure 4). From the assumption that g has degree 2, there are in fact two such connections (see also Lemma 2.3 below). This property, that there are two connections, and not only one, can cause resonance problems (on infinitely small scales) which complicates the analysis.

Since the location of the intersection between the two strips are central for the analysis, we define the set $I_0 \subset \mathbb{T}$ so that $I_0 + \omega$ is the projection onto the x -coordinate of the intersection between S and the strip $\mathbb{T} \times A'$ (we use the "thickened" interval A' instead of A to have some more space):

$$(2.6) \quad I_0 = \{x \in \mathbb{T} : (R + g(x)) \cap A' \neq \emptyset\} = \pi_1(S \cap (\mathbb{T} \times A')) - \omega.$$

From the above definitions, together with the assumptions on f and g , we get the following results:

Lemma 2.2. (1) If $(x, y) \in S$ and $x \notin I_0 + \omega$, then $y \in \mathbb{T} \setminus A'$. (2) If $(x, y) \in S$ and $x \notin \text{int}(I_0 + \omega)$, then $y \in \overline{\mathbb{T} \setminus A'} \subset B$.

Proof. This follows directly from the definition of S and (2.2). \square

Moreover, the set I_0 is small (see Figure 4):

Lemma 2.3. The set I_0 consists of two disjoint intervals, I_0^1, I_0^2 , each of a length $\sim \varepsilon$.

Proof. For completeness we include some details. By assumption we have $|R| < \varepsilon$, and we have $2\varepsilon < |A'| < 3\varepsilon$. Let $G(x)$ be a lift of $g(x)$. Since g has degree 2 we have $G(x+1) = G(x) + 2$. From the fact that the equation $g(x) = y$ has exactly two solutions $x_1, x_2 \in \mathbb{T}$ for each $y \in \mathbb{T}$, and since A' is a single interval and $|A'| \ll 1$, it is clear that I_0 consist of two intervals. Moreover, from the assumptions on g (condition (A1)(κ)) we have $\frac{1}{\kappa}(\xi - x) < G(\xi) - G(x) < \kappa(\xi - x)$ for all $x < \xi$. Thus, if $G(\xi) - G(x) = |A'| + |R|$ we have $(|A'| + |R|)/\kappa < \xi - x < \kappa(|A'| + |R|)$. This gives the bound on the length of each of the two intervals in I_0 . \square

The last result is an elementary fact which will be used at certain places later in the paper.

Lemma 2.4. If $I \subset \mathbb{T}$ is an interval such that $(I_0^1 + \omega) \subset I$, and $I \cap (I_0^2 + \omega) = \emptyset$, and if $\varphi : I \rightarrow \mathbb{T}$ is continuous and strictly increasing, and such that $(x, \varphi(x)) \in S$ for all $x \in I$, then $\varphi^{-1}(A')$ is a single interval, and $\varphi(\varphi^{-1}(A')) = A'$.

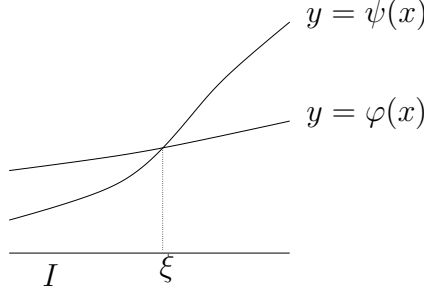
Proof. Since $(I_0^1 + \omega) \subset I$, and $(x, \varphi(x)) \in S$ for all $x \in I$, it follows from the definition of I_0 that $\varphi(I_0^1 + \omega) \supset A'$. Since φ is strictly increasing it thus follows that $\varphi^{-1}(A') \cap (I_0^1 + \omega)$ is a single interval. By the definition of I_0 we must have $\varphi^{-1}(A') \subset I_0 + \omega$. By assumption we have $I \cap (I_0^2 + \omega) = \emptyset$; thus $\varphi^{-1}(A') \subset I_0^1 + \omega$. See Figure 4. \square

2.4. Basic properties of the map T . Here we collect some basic information about iterations of the map T . We recall the notation

$$(x_k, y_k) = T^k(x, y), k \in \mathbb{Z}.$$

We begin with an obvious fact. However, since it is central for the analysis, we state it as a lemma.

Lemma 2.5. If $y, \eta \in \mathbb{T}$ and $x \in \mathbb{T}$, then $T(x, [y, \eta]) = \{x + \omega\} \times [y_1, \eta_1]$.

FIGURE 5. ψ overtakes φ at ξ .

Proof. Since f preserves the orientation we have $f([y, \eta]) = [y_1, \eta_1]$; here $y_1 = \pi_2(T(x, y))$ and $\eta_1 = \pi_2(T(x, \eta))$. \square

The statement of the next lemma was included in the discussion prior to the definition of the strip S ; but for later references we state it here as a lemma.

Lemma 2.6. *If $(x, y) \in \mathbb{T} \times (\mathbb{T} \setminus A)$ then $T(x, y) \in S$.*

Proof. Since $y \notin A$ it follows by definition that $f(y) \in R$; hence $g(x) + f(y) \in R + g(x)$. Thus $T(x, y) = (x + \omega, g(x) + f(y)) \in S$. \square

The previous lemma, combined with Lemma 2.2, gives us

Lemma 2.7. *If $(x, y) \in S$ and $x \notin \text{int}(I_0 + \omega)$ then $T(x, y) \in S$.*

The last lemma in this subsection deals with backward iteration of T .

Lemma 2.8. *If (x, y) are such that $x_1 \notin I_0 + \omega$ and $y_1 \in A'$, then $y \in A$.*

Proof. Assume that $y \in \mathbb{T} \setminus A$. Then $(x_1, y_1) \in S$ by Lemma 2.6. Since $y_1 \in A'$, it follows from the definition of I_0 that $x_1 \in I_0 + \omega$. Contradiction. \square

2.5. Some topological facts. Here we collect some elementary topological facts which we will use to control the geometry in our analysis. The lemmas below are tailored for the exact situation which we will have later.

First we introduce the following notation (see Figure 5).

Definition 2.9. Assume that $I = [s, t] \subset \mathbb{T}$ is an interval and that the functions $\varphi, \psi : I \rightarrow \mathbb{T}$ are continuous. Assume further that $\psi(\xi) = \varphi(\xi)$ for some $\xi \in I$, and that the lifts $\tilde{\varphi}, \tilde{\psi}$ are chosen so that $\tilde{\varphi}(\xi) = \tilde{\psi}(\xi)$. If $\tilde{\psi}(x) < \tilde{\varphi}(\xi)$ for all $x < \xi$ sufficiently close to ξ , and if $\tilde{\psi}(x) > \tilde{\varphi}(\xi)$ for all $x > \xi$ sufficiently close to ξ , we say that ψ overtakes φ at ξ .

An obvious case when this happens is when ψ is strictly increasing, and φ is a constant function. Then ψ overtakes φ at each point where the graphs intersect.

Since the map $f : \mathbb{T} \rightarrow \mathbb{T}$ preserves orientation, we immediately get:

Lemma 2.10. *Assume that $I = [s, t] \subset \mathbb{T}$ is an interval and that the functions $\varphi, \psi : I \rightarrow \mathbb{T}$ are continuous. Assume further that ψ overtakes φ at $x = \xi_i$ ($i = 1, 2, \dots, k$),*

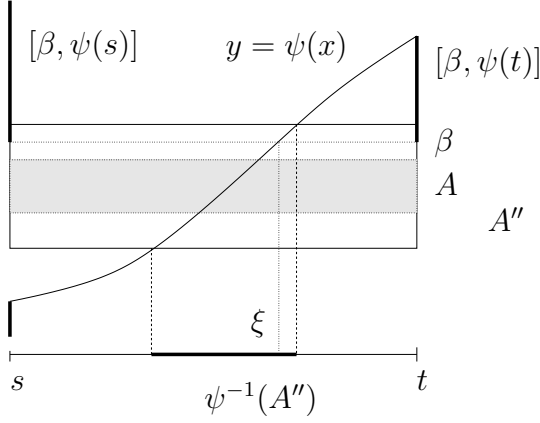


FIGURE 6. The statement of Lemma 2.11.

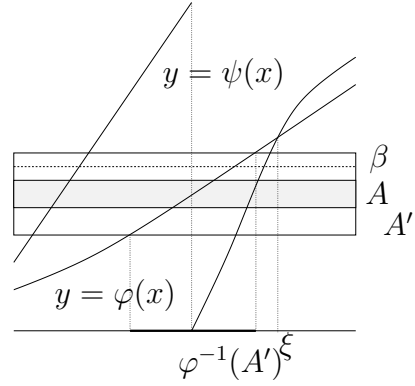


FIGURE 7. Figure for Lemma 2.12.

and that $\psi(x) \neq \varphi(x)$ for all $x \neq \xi_i$. Then the same holds for ψ_1 and φ_1 , where $\psi_1(x) = \pi_2(T(x, \psi(x)))$ and $\varphi_1(x) = \pi_2(T(x, \varphi(x)))$.

Proof. Note that $\psi_1(x) = g(x) + f(\psi(x))$ and $\varphi_1(x) = g(x) + f(\varphi(x))$, and recall that f is strictly increasing. \square

We recall the definition of the point β in (2.4) and that of $A'' \subset A'$ in (2.3).

Lemma 2.11. *Assume that $I = [s, t] \subset \mathbb{T}$ is an interval and that the function $\psi : I \rightarrow \mathbb{T}$ is continuous and strictly increasing. If $\psi^{-1}(A'')$ consists of one interval, I^1 , and $\psi(I^1) = A''$, then $[\beta, \psi(s)] \subset \mathbb{T} \setminus A$ and $[\beta, \psi(t)] \subset \mathbb{T} \setminus A$. Moreover, there is a unique $\xi \in I^1$ such that $\psi(\xi) = \beta$, this $\xi \in \text{int}(I^1)$ (and, trivially, ψ overtakes the constant function $\varphi(x) = \beta$ at ξ).*

Proof. Easy. See Figure 6 \square

Lemma 2.12. *Assume that $I = [s, t] \subset \mathbb{T}$ is an interval and that the functions $\varphi, \psi : I \rightarrow \mathbb{T}$ are continuous and strictly increasing, and that $\varphi^{-1}(A')$ consists of one interval, I^0 , and $\varphi(I^0) = A'$. Assume further that $[\varphi(s), \psi(s)], [\varphi(t), \psi(t)] \subset \mathbb{T} \setminus A$, and that there is a unique $\xi \in I$ such that $\varphi(\xi) = \psi(\xi)$, $\xi \in \text{int}(I)$, and that ψ overtakes φ at ξ . Then the equation $\psi(x) = \beta$ has exactly two solutions $\xi_1, \xi_2 \in I$, these $\xi_1, \xi_2 \in \text{int}(I)$ (and, trivially, ψ overtakes the constant function $\phi(x) = \beta$ at ξ_1, ξ_2). Furthermore, $[\beta, \psi(s)] \subset \mathbb{T} \setminus A$ and $[\beta, \psi(t)] \subset \mathbb{T} \setminus A$.*

Proof. We assume that $I = [s, t] \subset \mathbb{R}$. Let $[a'_1, a'_2] \subset \mathbb{R}$ be such that $\pi([a'_1, a'_2]) = A'$, and let $a'_1 < a_1 < a_2 < a'_2$ be such that $\pi((a_1, a_2)) = A$. We also let $a_2 < b < a'_2$ be such that $\pi(b) = \beta$ (recall the definition of β in (2.4)).

Let $\tilde{\varphi}$ be the lift of φ such that $a'_2 < \tilde{\varphi}(s) \leq a'_1 + 1$. From the assumptions we must have $a'_2 + 1 \leq \tilde{\varphi}(t) < a'_1 + 2$. We let $\tilde{\psi}$ be the lift of ψ such that $\tilde{\varphi}(s) < \tilde{\psi}(s) \leq a_1 + 1$ (recall that $[\varphi(s), \psi(s)] \subset \mathbb{T} \setminus A$). Note that $a_2 < b < \tilde{\psi}(s) \leq a_1 + 1$, and thus $[\beta, \psi(s)] \subset \mathbb{T} \setminus A$.

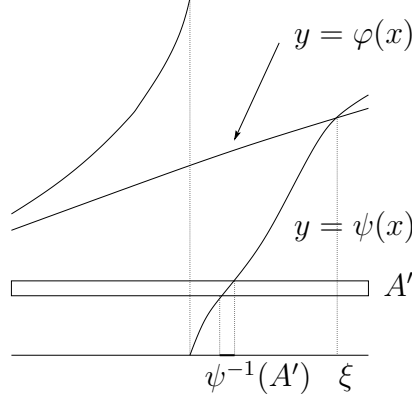


FIGURE 8. Figure for Lemma 2.13.

Since there is only one $\xi \in I$ such that $\varphi(\xi) = \psi(\xi)$, and since this ξ is in the interior of I , and ψ overtakes φ at ξ , we must have $\tilde{\psi}(\xi) = \tilde{\varphi}(\xi) + 1$, and $\tilde{\varphi}(x) + 1 < \tilde{\psi}(x) < \tilde{\varphi}(x) + 2$ for all $x > \xi$. The assumption that $[\varphi(t), \psi(t)] \subset \mathbb{T} \setminus A$ therefore gives us that $\tilde{\varphi}(t) + 1 < \tilde{\psi}(t) < a_1 + 3$. Since $b + 2 < a'_2 + 2 < \tilde{\varphi}(t) + 1$, this last inequality shows that $[\beta, \psi(t)] \subset \mathbb{T} \setminus A$. Moreover, we note that $[a_1 + 1, a'_2 + 2] \subset \tilde{\psi}(I) \subset (a'_2, a_1 + 3)$. Since $b + 1, b + 2 \in [a_1 + 1, a'_2 + 2]$, and $b, b + 3 \notin (a'_2, a_1 + 3)$, and since $\tilde{\psi}$ is strictly increasing, this shows that the equation $\psi(x) = \beta$ has exactly two solutions $\xi_1, \xi_2 \in I$, and that $\xi_1, \xi_2 \in \text{int}(I)$. \square

Lemma 2.13. *Assume that $I = [s, t] \subset \mathbb{T}$ is an interval and that the functions $\varphi, \psi : I \rightarrow \mathbb{T}$ are continuous and strictly increasing. Assume further that $\varphi(x) \in \mathbb{T} \setminus A'$ for all $x \in I$, and $[\varphi(s), \psi(s)], [\varphi(t), \psi(t)] \subset \mathbb{T} \setminus A'$.*

- (1) *If there is a unique $\xi \in I$ such that $\psi(\xi) = \varphi(\xi)$, if $\xi \in \text{int}(I)$, and if ψ overtakes φ at ξ , then $\psi^{-1}(A')$ consists of one single interval, I' , and $\psi(I') = A'$.*
- (2) *If there are exactly two solutions ξ_1, ξ_2 to $\psi(x) = \varphi(x)$, if $\xi_1, \xi_2 \in \text{int}(I)$, and if ψ overtakes φ at ξ_i ($i = 1, 2$), then $\psi^{-1}(A')$ consists of two interval, I', I'' , and $\psi(I') = \psi(I'') = A'$.*

Proof. We prove (1); the proof of (2) is similar. Let $[a_1, a_2] \subset \mathbb{R}$ be such that $\pi([a_1, a_2]) = A'$. By assumption we have $\varphi(I) \cap A' = \emptyset$. Assume that $I = [s, t] \subset \mathbb{R}$. Let $\tilde{\varphi}$ be the lift of φ such that $a_2 < \tilde{\varphi}(x) < a_1 + 1$ for all x , and let $\tilde{\psi}$ be the lift of ψ such that $\tilde{\varphi}(s) < \tilde{\psi}(s) < a_1 + 1$ (this is possible since $[\varphi(s), \psi(s)] \cap A' = \emptyset$). Since there is a unique $\xi \in I$ such that $\psi(\xi) = \varphi(\xi)$, this point being in the interior of I , and since ψ overtakes φ at ξ , we must have $\tilde{\varphi}(x) < \tilde{\psi}(x)$ for all $x \in I$. From this we conclude that we must have $\tilde{\psi}(\xi) = \tilde{\varphi}(\xi) + 1$, and $\tilde{\psi}(x) > \tilde{\varphi}(x) + 1$ for all $x > \xi$. Since $\tilde{\varphi}(t) + 1 < a_1 + 2$, since $[\varphi(t), \psi(t)] \cap A' = \emptyset$, and since $\tilde{\psi}(t) < \tilde{\varphi}(t) + 2$ (otherwise the equation $\psi(x) = \varphi(x)$ would have more than one solution), we see that we must have $\tilde{\psi}(t) < a_1 + 2$. From this the statement in (1) follows.

□

3. SOME FORMULAE

To control the dynamics of the map T we need to know how the iterates $(x_k, y_k) = T^k(x, y)$ depend on the initial point (x, y) . In this section we derive various formulae for this, as well as quantitative estimates.

We first collect a few well-known facts about Lipschitz functions.

3.1. Some properties of Lipschitz continuous functions. Assume that $\varphi, \psi : \mathbb{T} \rightarrow \mathbb{T}$ are Lipschitz continuous. Since Lipschitz continuity implies absolute continuity, it follows that φ and ψ are differentiable a.e.; and if $\tilde{\varphi}$ is a lift of φ , then the formula

$$\tilde{\varphi}(b) - \tilde{\varphi}(a) = \int_a^b \tilde{\varphi}'(x) dx$$

holds.

The sum $\varphi + \psi$ is clearly Lipschitz continuous. Moreover, the composition $\varphi \circ \psi$ is again Lipschitz continuous, and the chain rule holds, i.e.,

$$(\varphi \circ \psi)'(x) = \varphi'(\psi(x))\psi'(x) \text{ for a.e. } x \in \mathbb{T},$$

where $\varphi'(\psi(x))\psi'(x)$ is interpreted to be zero whenever $\psi'(x) = 0$.

If φ has a Lipschitz constant C , then $|\varphi'(x)| \leq C$ whenever the derivative exists; and if φ is bi-Lipschitz, where φ and φ^{-1} both have Lipschitz constant $C > 1$, then $C^{-1} \leq |\varphi'(x)| \leq C$.

We will sometimes use the following notation: if $\varphi : I \rightarrow \mathbb{T}$ is Lipschitz continuous, then " $|\varphi'(x)| \leq C$ on I " means that $|\varphi'(x)| \leq C$ whenever the derivative exists, i.e., for a.e. $x \in I$. Note that this implies, by the integral formula above, that φ has a Lipschitz constant C .

From the assumptions on f and g in (A1-2) we have the global bounds

$$\frac{1}{\kappa} < g'(x) < \kappa \text{ for a.e. } x \in \mathbb{T}, \quad \varepsilon^\rho < f'(y) < \varepsilon^{-\rho} \text{ for a.e. } y \in \mathbb{T};$$

and on the intervals $A \subset \mathbb{T}$ and $B \subset \mathbb{T}$ we have

$$f'(y) < \varepsilon \text{ on } B; \quad f'(y) > \varepsilon^{-1} \text{ on } A.$$

3.2. Formulae. For a fixed $x \in \mathbb{T}$ we consider the iterates $y_k : \mathbb{T} \rightarrow \mathbb{T}$ defined by

$$y_k(y) = \pi_2(T^k(x, y)).$$

Each of the iterates $y_1(y) = g(x) + f(y)$, $y_2(y) = g(x + \omega) + f(y_1(y))$, \dots are orientation preserving circle homeomorphisms, and they are all Lipschitz continuous by the properties in the previous subsections. We have $y'_1(y) = f'(y)$, $y'_2(y) = f'(y_1)f'(y)$, and generally

$$y'_{k+1}(y) = \prod_{j=0}^k f'(y_j(y)) \text{ for a.e. } y \in \mathbb{T}.$$

Furthermore, if $Y \subset \mathbb{T}$ is an interval, then

$$|y_k(Y)| = \int_Y y'_k(y) dy.$$

Below we shall often keep track on products (for fixed x) of the form $\prod_{j=0}^k f'(y_j(y))$ for a.e. y in some interval Y ; and our formulae will contain such expressions. We note that $y_{k+1}(y)$ having a Lipschitz constant C on Y is equivalent to $\prod_{j=0}^k f'(y_j(y)) \leq C$ on Y since we know that y_{k+1} is Lipschitz continuous (and $f'(y)$ is always positive).

From these observations we immediately get

Lemma 3.1. *Assume that $Y \subset \mathbb{T}$ is an interval, and that for some $k \geq 1$ and $x \in \mathbb{T}$*

$$\prod_{j=0}^{k-1} f'(y_j(y)) \underset{(\geq)}{\leq} \delta \quad \text{on } Y.$$

Then $|y_k(Y)| \underset{(\geq)}{\leq} \delta |Y|$.

We shall need formulae to control the iterates (under T) of curves $(x, y_0(x))$, $x \in I$, where I is an interval.

Lemma 3.2. *Assume that $y_0 : I \rightarrow \mathbb{T}$ (I an interval) is Lipschitz continuous, and define the functions $y_j(x)$ by $y_j(x) = \pi_2(T^j(x, y_0(x)))$. For each $k \geq 0$ the function y_{k+1} is Lipschitz continuous on I and we have the formula*

$$\begin{aligned} y'_{k+1}(x) &= g'(x_k) + g'(x_{k-1})f'(y_k) + g'(x_{k-2}) \prod_{j=k-1}^k f'(y_j) + \dots + g'(x) \prod_{j=1}^k f'(y_j) + \\ &+ y'_0(x) \prod_{j=0}^k f'(y_j) \text{ for a.e. } x \in I. \end{aligned}$$

Proof. Note that $y_1(x) = g(x) + f(y_0(x))$; thus $y'_1(x) = g'(x) + f'(y_0)y'_0(x)$. Next, $y_2(x) = g(x_1) + f(y_1(x))$, and therefore $y'_2(x) = g'(x_1) + f'(y_1)y'_1(x) = g'(x_1) + f'(y_1)(g'(x) + f'(y_0)y'_0(x))$. The general formula follows by induction. \square

As a direct consequence of the previous lemma we get the following result, where we assume that we for each $x \in I$ have various controls on products of the form

$$\prod_{j=\ell}^k f'(y_j)$$

where, as usual, $y_j = \pi_2(T^j(x, y))$.

Lemma 3.3. *Below $I \subset \mathbb{T}$ and $Y \subset \mathbb{T}$ are assumed to be intervals.*

- (1) Assume that $0 < \delta < 1$ and $k \geq 1$. Assume further that we for all $x \in I$ have the estimate $\prod_{j=\ell}^k f'(y_j) < \delta^{k-\ell+1}$ on Y for all $\ell \in [0, k]$. If $y_0 : I \rightarrow Y$ is Lipschitz continuous and increasing, then

$$y'_{k+1}(x) < \frac{\kappa}{1-\delta} + \delta^{k+1}y'_0(x) \text{ on } I.$$

- (2) Assume that $\Delta > 0$. Moreover, assume that for all $x \in I$ we have $\prod_{j=0}^k f'(y_j) > \Delta$ for all y such that $y_{k+1} \in Y$ (and for which the derivatives exist). If $y_0 : I \rightarrow \mathbb{T}$ is Lipschitz continuous and increasing, and if $y_{k+1}(I) \subset Y$, then

$$y'_{k+1}(x) > \Delta y'_0(x) \text{ on } I.$$

- (3) For all increasing Lipschitz functions $y_0 : \mathbb{T} \rightarrow \mathbb{T}$ we always have the (trivial) upper bound

$$y'_{k+1}(x) < \frac{\kappa \varepsilon^{-\rho k}}{1 - \varepsilon^\rho} + \varepsilon^{-\rho(k+1)} y'_0(x) \text{ on } I, k \geq 0.$$

- (4) For all increasing Lipschitz functions $y_0 : \mathbb{T} \rightarrow \mathbb{T}$ always have the (trivial) lower bound $y'_{k+1}(x) > 1/\kappa$ on I , $k \geq 0$.

Proof. (1) From the formula in Lemma 3.2, combined with the assumed estimates and the general assumption that $g'(x) < \kappa$ for a.e. x , we get

$$y'_{k+1} < \kappa(1 + \delta + \delta^2 + \dots + \delta^k) + \delta^{k+1}y'_0 < \frac{\kappa}{1-\delta} + \delta^{k+1}y'_0.$$

- (2) Since all the terms in the formula in Lemma 3.2 are positive, we get, making use of the assumption,

$$y'_{k+1} > \Delta y'_0.$$

- (3) Follows easily from the formula in Lemma 3.2 since we have the global upper bounds $f'(y) < \varepsilon^{-\rho}$ and $g'(x) < \kappa$.

- (4) Again, since all the terms in the formula are non-negative, and since $g'(x) > 1/\kappa$ for a.e. x , the statement is obvious. \square

We end this section with an elementary result which we will use later to show that certain sets are "well" inside other sets. We recall the definition of the set A'' in (2.3).

Lemma 3.4. *Assume that $I = [s, t] \subset \mathbb{T}$ is an interval and that $\varphi : I \rightarrow \mathbb{T}$ is Lipschitz continuous, strictly increasing, and such that $\varphi'(x) < \Delta$ on I . Assume further that $\varphi(s), \varphi(t) \notin A''$. Then $\text{dist}(\varphi^{-1}(A), \partial I) > \varepsilon/(2\Delta)$.*

Proof. Let $\tilde{\varphi} : \mathbb{R} \rightarrow \mathbb{R}$ be a lift of φ , and assume that $I = [s, t] \subset \mathbb{R}$. Let $s < s'$ be the smallest s' such that $\pi(\tilde{\varphi}(s')) \in \overline{A}$ (we assume that $\varphi^{-1}(A) \neq \emptyset$). From the assumption that $\varphi(s) \notin A''$, it follows that $\tilde{\varphi}(s') - \tilde{\varphi}(s) > \varepsilon/2$. Moreover, $\tilde{\varphi}(s') - \tilde{\varphi}(s) =$

$\int_s^{s'} \tilde{\varphi}'(x)dx \leq \Delta(s' - s)$. Thus, $\text{dist}(\varphi^{-1}(A), s) > \varepsilon/(2\Delta)$. In the same way we can prove that $\text{dist}(\varphi^{-1}(A), t) > \varepsilon/(2\Delta)$. \square

4. PROOF OF THEOREM 1 (ASSUMING GOOD CONTROL ON THE DYNAMICS OF T)

In this section we prove Theorem 1, assuming that we have very good control of the dynamics of the map T . The estimates we will use are contained in the next proposition, Proposition 4.1. The proof of this proposition is the content of sections 5 and 6 below.

Before stating the proposition we recall the notation: given $(x, y) \in \mathbb{T}^2$ we define (x_n, y_n) by

$$(x_n, y_n) = T^n(x, y), \quad n \in \mathbb{Z}.$$

We also recall that we have fixed ω satisfying the Diophantine condition (1.3) for some $\gamma > 0$, and that we have assumed that g satisfies condition (A1)(κ) and f satisfies (A2)(ε, ρ). Finally, we recall that $A \subset A'$ (see the definition of A' in (2.1)).

Proposition 4.1. *There exists an $\varepsilon_1 = \varepsilon_1(\gamma, \kappa, \rho) > 0$ such that for every $0 < \varepsilon < \varepsilon_1$ there is a measurable set $X \subset \mathbb{T}$, $|X| > 0$, such that the following holds for each $x \in X$:*

(1) *For a.e. $y \in \mathbb{T} \setminus A$ we have*

$$\prod_{j=1}^k f'(y_j) < \varepsilon^{k/2} \text{ for all } k \geq 1.$$

(2) *For all $y \in \mathbb{T} \setminus A$ we have that whenever $x_j \in X$ ($j \geq 1$), then $y_j \in \mathbb{T} \setminus A'$. Moreover, $\overline{\{(x_k, y_k)\}_{k \geq 0}} = \mathbb{T}^2$.*

(3) *If $Y \subset \mathbb{T}$ is an interval such that $(x_k, y_k) \in X \times A'$ for all $y \in Y$, any $k \geq 1$, then*

$$\prod_{j=0}^{k-1} f'(y_j) > \varepsilon^{-k/2} \text{ for a.e. } y \in Y.$$

Remark 5. We could have replaced condition (3) with a condition similar to that of (1), but instead considering backward iterations. However, since we do not really need such a condition we have chosen to work with the above form.

Proof of Theorem 1. We shall first construct two measurable functions $u : \mathbb{T} \rightarrow \mathbb{T}$ and $s : \mathbb{T} \rightarrow \mathbb{T}$, whose graphs are a.e. T -invariant. Moreover, we will see that for a.e. $x \in \mathbb{T}$ we have

$$d(y_k, u(x_k)) \rightarrow 0 \text{ (exponentially fast) as } k \rightarrow \infty \text{ for all } y \neq s(x).$$

(u, s are the equivalent of the (projective) Oseledets' directions in the linear cocycle case. See [21].)

Let

$$X^* = \{x \in \mathbb{T} : x + j\omega \in X \text{ for infinitely many } j \geq 0 \text{ and infinitely many } j < 0\}.$$

Since $|X| > 0$ it follows immediately from the Poincaré recurrence theorem that X^* has full (Lebesgue) measure; and obviously X^* is invariant under rotation by ω . We now construct $u(x), s(x)$ for each $x \in X^*$.

Fix $x \in X^*$ and let $\xi = x_p$, where $p \geq 0$ is the smallest integer such that $x + p\omega \in X$. We shall now focus on iterates of (ξ, y) for $y \in \mathbb{T}$. Let $0 = j_0 < j_1 < j_2 < \dots$ be the times when $\xi_{j_i} \in X$.

Claim 1: If y is such that $y_{j_i} \in A'$ ($i \geq 1$), then $y_{j_l} \in A$ for all $0 \leq l < i$.

Indeed, if $y_{j_l} \notin A'$ for some $0 \leq l < j$, then, since $\xi_{j_l} \in X$, applying Proposition 4.1(2) to (ξ_{j_l}, y_{j_l}) would give us that $y_{j_l} \notin A'$.

Define the closed intervals $S_i = \pi_2(T^{-j_i}(\xi_{j_i}, A'))$, $i \geq 1$. From Claim 1 it follows that $S_{i+1} \subset S_i \subset A$ for all $i \geq 1$. Moreover, from Proposition 4.1(3) combined with Lemma 3.1 it follows that $|S_i| \rightarrow 0$ as $i \rightarrow \infty$. Let $s(\xi)$ be the unique point in $\bigcap_{i \geq 1} S_i$. We now define $s(x)$ by $s(x) = \pi_2(T^{-p}(\xi, s(\xi)))$. Note that, since $s(\xi) \in S_i$ for each $i \geq 1$ (by construction), we have $T^{p+j_i}(x, s(x)) = T^{j_i}(\xi, s(\xi)) \in X \times A'$ for all $i \geq 1$. Using Claim 1 we conclude that we in fact have

$$(4.1) \quad T^{p+j_i}(x, s(x)) = T^{j_i}(\xi, s(\xi)) \in X \times A \text{ for all } i \geq 1.$$

To construct $u(x)$ we do as follows. Let $\theta = x_{-q}$, where $q \geq 0$ is the smallest integer such that $x - q\omega \in X$. Let $0 > k_1 > k_2 > \dots$ be the (negative) times when $\theta_{k_i} \in X$, and define the closed intervals $U_i = \pi_2(T^{-k_i}(\theta_{k_i}, \mathbb{T} \setminus A))$, $i = 1, 2, \dots$. Proposition 4.1(2) implies that $U_{i+1} \subset U_i \subset \mathbb{T} \setminus A'$. Furthermore, Proposition 4.1(1), and Lemma 3.1, give that $|U_i| \rightarrow 0$ as $i \rightarrow \infty$. Let $u(\theta)$ be the unique point in $\bigcap_{i \geq 1} U_i$, and define $u(x) = \pi_2(T^q(\theta, u(\theta)))$. Since $(\theta, u(\theta)) \in X \times (\mathbb{T} \setminus A)$, and since $x = \theta + q\omega$, it follows from Proposition 4.1(2) that

$$(4.2) \quad T^{p+j_i}(x, u(x)) = T^{j_i+q+p}(\theta, u(\theta)) \in X \times (\mathbb{T} \setminus A') \text{ for each } i \geq 1.$$

By combining equations (4.1) and (4.2) it follows that $s(x) \neq u(x)$ for all $x \in X^*$. Moreover, from the above construction of $u(x)$ and $s(x)$ it is clear that $T(x, u(x)) = (x + \omega, u(x + \omega))$ and $T(x, s(x)) = (x + \omega, s(x + \omega))$ for all $x \in X^*$.

Next we show that

$$d(y_n, u(x_n)) \rightarrow 0 \text{ as } n \rightarrow \infty \text{ for all } x \in X^*, y \neq s(x).$$

To do this we first note that Proposition 4.1(1) and Lemma 3.1 imply that

$$(4.3) \quad |\pi_2(T^n(z, \mathbb{T} \setminus A))| \rightarrow 0 \text{ (exponentially fast) as } n \rightarrow \infty \text{ for all } z \in X.$$

Fix $x \in X^*$, and let p, ξ, j_i, S_i be defined as above. Take $y \neq s(x)$. If $y_{p+j_i} \in \mathbb{T} \setminus A'$ for some i , then it follows from (4.2) and (4.3) that $d(y_n, u(x_n)) \rightarrow 0$ as $n \rightarrow \infty$. Assume now that $y_{p+j_i} \in A'$ for all $i \geq 1$. Then, by definition, $y_p \in S_i$ for all $i \geq 1$. But this means that $y_p \in \bigcap_{i \geq 1} S_i$, i.e., $y_p = s(\xi)$; hence $y = s(x)$, contradicting the assumption.

It remains to show that T is minimal. By Proposition 4.1(2) we have

$$(4.4) \quad \overline{(x_k, y_k)_{k \geq 0}} = \mathbb{T}^2 \text{ for all } x \in X, y \in \mathbb{T} \setminus A.$$

The plan is to use Proposition 7.5 in Section 7 with $w^+ = u$ and $w^- = s$ (note that all the conditions needed for applying the proposition are fulfilled). From that proposition we get

that T has exactly two invariant and ergodic Borel probability measures, μ^s, μ^u . Moreover, μ^u and μ^s are the push-forward of the Lebesgue measure on \mathbb{T} by the maps $x \mapsto (x, u(x))$ and $x \mapsto (x, s(x))$, respectively. Proposition 7.5 also tells us that $\text{supp } \mu^u$ is a minimal set. From (4.4) and (4.2) we conclude that $\bigcup_{k=0}^{\infty} \{(x + k\omega, u(x + k\omega))\} = \mathbb{T}^2$ for all $x \in X^*$. Thus $\text{supp } \mu^u = \mathbb{T}^2$, i.e., \mathbb{T}^2 is a minimal set. \square

The remaining part of the paper is now devoted to the proof of Proposition 7.5. The set $X \subset \mathbb{T}$ in that proposition will be constructed inductively. We will obtain sets

$$X_0 \supset X_1 \supset \dots \supset X_n \supset \dots \supset X$$

where we get control on the iteration of points (x, y) , $x \in X_n$, for times $T_n \rightarrow \infty$. The Diophantine condition (1.3) posed on ω is crucial for the analysis; it is used to obtain stopping times.

5. BASE CASE OF THE CONSTRUCTION

We are now ready to take the first steps towards the proof of Proposition 4.1. The proof of this proposition is the heart of our construction, which is based on an inductive argument.

In this section we state and prove the base case for the inductive construction. Several of the geometric ideas used in the general inductive step are visible already here. The main approach is similar to the one we used in [4], but the geometry is very different (due to the fact that the map T is not homotopic to the identity). This is why, in particular, we do not need much regularity assumptions on f and g .

5.1. Control on iterations on first scale. From the basic estimates in Section 2.4, we immediately get good control on the iterates of points $(x, y) \in \mathbb{T} \times (\mathbb{T} \setminus A)$ up to time $N(x; I_0)$ (recall that this is the first entry time of the point x to I_0 under translation by ω). We summarize them in the next lemma.

Lemma 5.1. *The following holds:*

- (1) *Assume that $x \in \mathbb{T}$, and let $N = N(x; I_0)$. For all $y \in \mathbb{T} \setminus A$ we have $y_j \in \mathbb{T} \setminus A'$ for all $j \in [1, N]$ and $(x_j, y_j) \in S$ for all $j \in [1, N + 1]$. In particular, for all $1 \leq \ell \leq k \leq N$ we have*

$$\prod_{j=\ell}^k f'(y_j) < \varepsilon^{(k-\ell+1)} \text{ on } \mathbb{T} \setminus A.$$

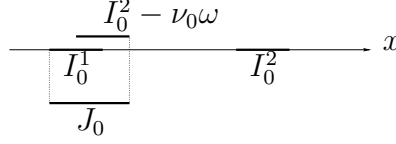
- (2) *Given $k \geq 1$, assume that x is such that $x \in \left(\mathbb{T} \setminus \bigcup_{j=0}^{k-1} (I_0 - j\omega)\right)$ and y is such that $y_k \in A'$. Then*

$$y_j \in A, \quad j = 0, \dots, k-1,$$

and thus

$$\prod_{j=0}^{k-1} f'(y_j) > \varepsilon^{-k}$$

whenever the derivatives exist.

FIGURE 9. Definition of J_0 in the resonant case.

Proof. (1) By Lemma 2.6 we have $(x_1, y_1) \in S$. The results in the statement of (1) now follows by a repeated use of Lemmas 2.2 and 2.7, and by recalling that $f'(\eta) < \varepsilon$ for a.e. $\eta \in B$.

(2) From the assumption on x it follows that $x_j = x + j\omega \notin I_0 + \omega$ for all $j = 1, 2, \dots, k$. Since $A \subset A'$, and since $f'(\eta) > 1/\varepsilon$ for a.e. $\eta \in A$, the result follows by applying Lemma 2.8 repeatedly. \square

To obtain control of orbits for longer time than $N(x; I_0)$ we have to analyse what can happen after we have enter I_0 , and what is the probability for the different scenarios. To do this we will send in "probes" through I_0 which we then can follow when we iterate points. However, the first step is to check if the two intervals in I_0 are what we call "resonant" or not.

5.2. Definition of the sets J_0, \widehat{J}_0 and the integers M_0, K_0 . We know (Lemma 2.3) that the set I_0 consists of two intervals, I_0^1 and I_0^2 , each of length $\sim \varepsilon$. Now it is time to use the Diophantine condition on ω . By applying Lemma 7.1 we see that

$$(5.1) \quad I_0^i \cap (I_0^i + k\omega) = \emptyset \quad (i = 1, 2)$$

for all $0 < |k| < \sqrt{\gamma/|I_0^i|} \sim \varepsilon^{-1/2}$; i.e, if we start in I_0^i it takes many steps (under rotation by ω) before we enter I_0^i again. However, if we, for example, start in I_0^1 , we may very well enter I_0^2 in just a few steps. We therefore need to check if the intervals I_0^1, I_0^2 are what we call "resonant". By this we mean that one of them is close to a translation (by few steps) of the other. We shall now fix different scales, depending on whether the intervals I_0^1 and I_0^2 are resonant or not. We stress that we do not seek for optimal choices, in the sense that the $\varepsilon_0(\gamma, \kappa, \rho)$ in Proposition 4.1 is as small as possible. Any such ε_0 is enough for us.

Case R: If there is a $\nu_0, 0 < \nu_0 \leq \varepsilon^{-1/40}$ such that

$$I_0^1 \cap (I_0^2 - \nu_0\omega) \neq \emptyset,$$

(if $I_0^2 \cap (I_0^1 - \nu_0\omega) \neq \emptyset$ we relabel the intervals) let J_0 be the single interval

$$J_0 = I_0^1 \cup (I_0^2 - \nu_0\omega)$$

and let

$$K_0 = [\varepsilon^{-1/20}] \text{ and } M_0 = K_0^2 = [\varepsilon^{-1/20}]^2.$$

By applying Lemma 7.2, we see that

$$(5.2) \quad \begin{aligned} (J_0 + k\omega) \cap I_0 &= \emptyset \text{ for all } k \in [-M_0^2, M_0^2] \setminus \{0, \nu_0\}, \text{ and} \\ J_0 \cap I_0 &= I_0^1, \quad (J_0 + \nu_0\omega) \cap I_0 = I_0^2. \end{aligned}$$

We note that $0 < \nu_0^2 \leq K_0$. We also note, by applying the same argument as in (5.1), using the fact that J_0 is a single interval and $|3J_0| \sim \varepsilon$, that

$$(5.3) \quad 3J_0 \cap (3J_0 + \omega) = \emptyset \text{ for all } 0 < |k| \leq M_0^2.$$

Case NR: Otherwise, i.e., if $I_0 \cap (I_0 + k\omega) = \emptyset$ for all $0 < |k| \leq \varepsilon^{-1/40}$, let $J_0 = I_0$ (thus J_0 consists of two intervals), and

$$K_0 = \lceil \varepsilon^{-1/160} \rceil \text{ and } M_0 = K_0^2 = \lceil \varepsilon^{-1/160} \rceil^2.$$

We note that $M_0^2 \leq \varepsilon^{-1/40}$. We thus have

$$(5.4) \quad J_0 \cap (J_0 + k\omega) = \emptyset \text{ for all } 0 < |k| \leq M_0^2.$$

Since each interval in J_0 is of length $\sim \varepsilon$ we also have

$$(5.5) \quad 3J_0^i \cap (3J_0^i + k\omega) = \emptyset \text{ for all } 0 < |k| \leq M_0^2 \text{ and } i = 1, 2.$$

For a (small) technical reason (at one place we would have liked $\mathbb{T} \setminus J_0$ to be closed) we now define an open set $\widehat{J}_0 \supset J_0$; we just make J_0 slightly bigger. In Case R we let \widehat{J}_0 be a single open interval containing J_0 , and so close to J_0 that (5.3) holds with J_0 replaced by \widehat{J}_0 (since J_0 is closed, (5.3) is an open condition). We also assume that $|\widehat{J}_0 \setminus J_0| < \varepsilon$. In Case NR we cover each interval J_0^i with an open interval \widehat{J}_0^i , close to J_0^i , such that (5.5) holds with $3J_0^i$ replaced by $3\widehat{J}_0^i$, and such that (5.4) holds with J_0 replaced by $\widehat{J}_0 = \widehat{J}_0^1 \cup \widehat{J}_0^2$. Finally we assume that $|\widehat{J}_0^i \setminus J_0^i| < \varepsilon$ ($i = 1, 2$).

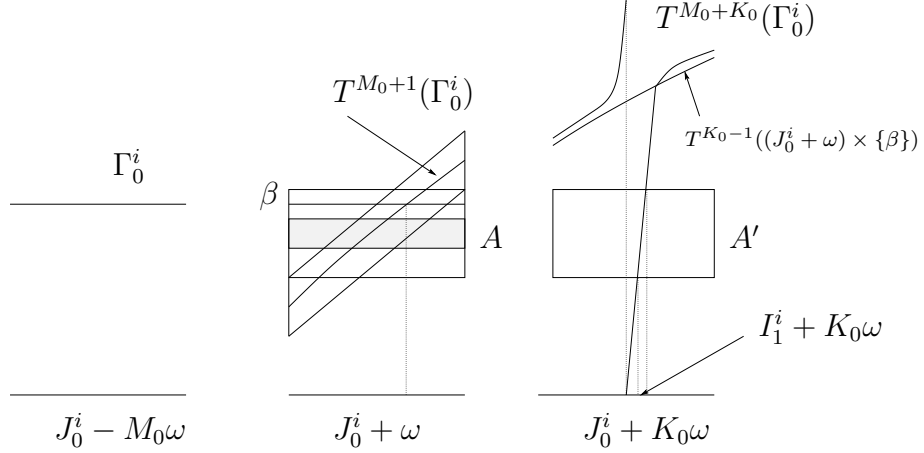
The purpose of the M_0 and K_0 is the following. We want to focus on points $x \in \mathbb{T}$ which do not enter (under translation) the "bad" set I_0 too fast (for at least M_0 steps). If $y \in \mathbb{T} \setminus A$ we can then use Lemma 5.1 to get a good control on the iterates; and $\prod_{k=1}^{N(x; I_0)} f'(y_j)$ is very small. We cannot avoid entering I_0 , and once we do that we lose control on the further iterates. However, we will see that there is only for a very tiny set of x (which we started with) for which we have not regained control after K_0 steps from when we entered I_0 ; and once we have control again, we know that we will not enter the bad set I_0 for a long time ($\gg M_0$). The situation is complicated by the fact that I_0 consists of two intervals; thus we may not have recovered from the first "hit" before we enter I_0 again. This is why we introduced the set J_0 , in order to keep track on when we have "fast returns" (resonances) or not.

From the above definitions we see that $N(x; J_0) \leq N(x; I_0)$ for "many" x . Namely,

Lemma 5.2. *If $x \in X_0 = \mathbb{T} \setminus \bigcup_{m=-M_0+1}^{K_0-1} (J_0 + m\omega)$, then $N(x; J_0) \leq N(x; I_0)$.*

Remark 6. Note that $|\bigcup_{m=-M_0+1}^{K_0-1} (J_0 + m\omega)| \leq (M_0 + K_0)|J_0| \ll \sqrt{\varepsilon}$, so X_0 is a large set.

Proof. In the non-resonant case, when $I_0 = J_0$, there is nothing to prove. Assume that we are in the resonant case, and let $m = N(x; I_0)$. Since $J_0 = I_0^1 \cup (I_0^2 - \nu_0\omega)$, and $0 < \nu_0 \ll K_0$, it follows from the assumption on x that $m > \nu_0$. If $x_m \in I_0^1 \subset J_0$, then $N(x; J_0) \leq m$; if $x_m \in I_0^2 \subset J_0 + \nu_0\omega$, then $x_{m-\nu_0} \in J_0$, and thus $N(x; J_0) \leq m - \nu_0$. \square

FIGURE 10. The geometry in the non-resonant case ($i = 1, 2$).

5.3. Geometry control. Now we will send in a "probe" through J_0 , which we later can follow ("most" iterates will cluster around the probe) in the transition from J_0 to $J_0 + K_0\omega$. The next lemma contains the geometric information about this probe. Obtaining control on the geometry, both in the base case as well as in the inductive step, is the trickiest part in the proof of Proposition 4.1. However, the topological properties of the map T helps quite a bit (some of the properties are formulated in the lemmas in Section 2.5).

Recall that the reference point $\beta \in \mathbb{T} \setminus A$ was defined in (2.4). At many occasions in the paper we will use the crucial fact that if the functions $y_k(x)$ are defined by $y_k(x) = \pi(T^k(x, \beta))$, then, by Lemma 3.3(4), we have $y'_k(x) > 1/\kappa$ on \mathbb{T} for all $k \geq 1$. In particular, the functions are strictly increasing.

Lemma 5.3. *Let $\Gamma_0 = (J_0 - M_0\omega) \times \{\beta\}$, and define $I_1 \subset J_0$ by*

$$I_1 = \pi_1(T^{M_0+K_0}(\Gamma_0) \cap (\mathbb{T} \times A')) - K_0\omega.$$

Then I_1 consists of two disjoint intervals, I_1^1 and I_1^2 , and $\text{dist}(I_1, \partial J_0) > 0$. Moreover, if we write $T^{M_0+K_0}(\Gamma_0) = \{(x + K_0\omega, \varphi_0(x)) : x \in J_0\}$, then $\varphi(I_1^i) = A'$ and $\varepsilon^{-K_0/2} < \varphi'_0(x) < \varepsilon^{-\rho K_0}$ on I_1^i ($i = 1, 2$).

Proof. By the definition of Γ_0 we have

$$\varphi_0(x) = \pi_2(T^{M_0+K_0}(x - M_0\omega, \beta)).$$

We divide the analysis into two cases, depending on whether J_0 consists of one or two intervals (recall the definition of J_0 in the previous section). (If g would have degree 1, and not 2 as we have assumed, the set I_0 would only contain one single interval, and we would only have the first of the two cases below; no resonances would be possible.)

Case NR ($J_0 = I_0$; the non-resonant case). We focus on one of the two intervals in I_0 , call it I_0^1 . The other interval is treated in exactly the same way. We shall write $I_0^1 = [s, t]$.

Let

$$y_k(x) = \pi_2(T^k(x - M_0\omega, \beta)).$$

Then we can write $T^k(\Gamma_0) = \{(x + (k - M_0\omega), y_k(x)) : x \in I_0\}$. Thus, to describe the part of the set I_1 which lies in I_0^1 , we need to control $y_{K_0+M_0}(x)$ for $x \in I_0^1$. We note that if $\theta \in I_0 - M_0\omega$, then $N(\theta; I_0) = M_0$ (by (5.4)). Since $\beta \in \mathbb{T} \setminus A$, it therefore follows from Lemma 5.1(1) that

$$(5.6) \quad (x + \omega, y_{M_0+1}(x)) \in S, \quad x \in I_0^1.$$

Moreover, it also follows that for all $x \in I_0 - M_0\omega$ we get

$$\prod_{j=\ell}^N f'(\sigma_j) < \varepsilon^{N-\ell+1} \text{ for all } 1 \leq \ell \leq N \text{ and a.e. } \sigma \in \mathbb{T} \setminus A$$

(here $\sigma_j = \pi_2(T^j(x, \sigma))$). By applying Lemma 3.3(1), and using the trivial lower bound from Lemma 3.3(4) and the fact that y_0 is constant, we get

$$(5.7) \quad \frac{1}{\kappa} < y'_{M_0+1}(x) < 2\kappa \text{ on } I_0^1.$$

From Lemma 3.3(3) it therefore follows that we have the upper bound

$$y'_{M_0+K_0}(x) < 3\kappa\varepsilon^{-\rho(K_0-1)} < \varepsilon^{-\rho K_0} \text{ on } I_0^1.$$

Combining (5.6) with Lemma 2.4 we see that

$$(5.8) \quad \{x \in I_0^1 : y_{M_0+1}(x) \in A'\} \text{ is a single interval, and } y_{M_0+1}(I_0^1) \supset A'.$$

Next we show that the part of I_1 in I_0^1 is well-inside I_0^1 , namely that

$$\text{dist}(I_1 \cap I_0^1, I_0^1) > \varepsilon/(4\kappa).$$

To do this we note that $N(x; I_0) \gg K_0$ for $x \in I_0 + \omega$, and since $(I_0 + K_0\omega) \cap (I_0 + \omega) = \emptyset$ (by (5.4)), it follows from Lemma 5.1(1) that $I_1 \cap I_0^1 \subset I_1' := \{x \in I_0^1 : y_{M_0+1}(x) \in A'\}$. (Indeed, if $(\theta, \eta) \in (I_0^1 + \omega) \times (\mathbb{T} \setminus A)$, then, by Lemma 5.1(1), we get $\eta_{K_0-1} \notin A'$.) Hence it follows from Lemma 3.4 and the bound in (5.7) that $\text{dist}(I_1 \cap I_0^1, I_0^1) > \varepsilon/(4\kappa)$.

To get control on $\varphi_0(x) = y_{M_0+K_0}(x)$ for $x \in I_0^1$, and thus control one of the two pieces of Γ_0 , we will "shadow" the orbit of $(x + \omega, y_{M_0+1}(x))$, for $x \in I_0^1$, with another one which is easier to follow. To do this, we let

$$\eta_k(x) = \pi_2(T^k(x + \omega, \beta)).$$

By applying Lemma 2.11 to the restriction of $y_{M_0+1}(x)$ to $I_0^1 = [s, t]$, we see that (recall (5.8) and note that $\eta_0 = \beta$)

$$(5.9) \quad [\eta_0, y_{M_0+1}(s)], [\eta_0, y_{M_0+1}(t)] \subset \mathbb{T} \setminus A,$$

and there is a unique $\xi \in I_0^1$ such that $y_{M_0+1}(\xi) = \eta_0$, this $\xi \in \text{int}(I_0^1)$, and y_{M_0+1} overtakes η_0 at ξ . See Fig. 10. By Lemma 2.10 we thus know that y_{M_0+1+k} overtakes η_k at ξ for all k (and ξ obviously is the unique point of intersection in I_0^1). Since $N(x + \omega; I_0) \gg K_0$ for all $x \in I_0$ (by (5.4)), and since $\beta \in \mathbb{T} \setminus A$, it follows from Lemma 5.1(1) that

$$\eta_{K_0-1}(x) \in \mathbb{T} \setminus A' \text{ for all } x \in I_0^1;$$

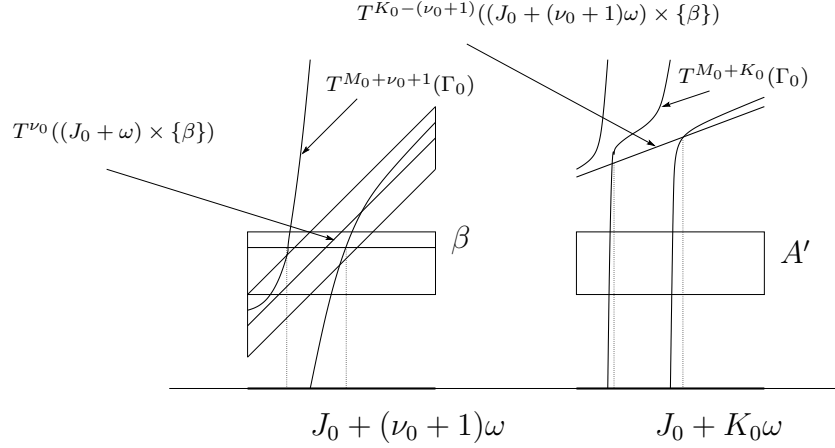


FIGURE 11. The additions in the resonant case.

and since (5.9) holds, the same argument shows that (recall Lemma 2.5(1))

$$[\eta_{K_0-1}(s), y_{M_0+K_0}(s)], [\eta_{K_0-1}(t), y_{M_0+K_0}(t)] \subset \mathbb{T} \setminus A'.$$

We are thus in a situation where we can apply Lemma 2.13(1) to conclude that there is a single interval $I_1^1 \subset I_0^1$ such that $I_1^1 = \{x \in I_0^1 : y_{M_0+K_0}(x) \in A'\}$, and $y_{M_0+K_0}(I_1^1) = A'$.

It remains to check that $y'_{M_0+K_0}(x)$ is large on I_1^1 . To do this we note that, by (5.4), we have $I_0 + \omega \subset \mathbb{T} \setminus \bigcup_{j=0}^{K_0-1} (I_0 - j\omega)$. It therefore follows from Lemma 5.1(2) that if $(\theta, \sigma) \in (I_0 + \omega) \times \mathbb{T}$ are such that $\sigma_{K_0-2} \in A'$, then

$$\prod_{j=0}^{K_0-2} f'(\sigma_j) > \varepsilon^{-(K_0-1)}$$

whenever the derivatives exist. Since $T^{M_0+1}(\Gamma_0) = \{(x + \omega), y_{M_0+1}(x) : x \in I_0\}$, it thus follows from Lemma 3.3(2), together with the estimates on y_{M_0+1} in (5.7) and the definition of I_1^1 , that $y'_{M_0+K_0}(x) > \varepsilon^{-(K_0-1)}/\kappa \gg \varepsilon^{-K_0/2}$ for a.e. $x \in I_1^1$.

Case R ($J_0 = I_0^1 \cup (I_0^2 - \nu_0\omega)$; *resonant case*). We shall write the single interval $J_0 = [s, t]$. As in the previous case we let

$$y_k(x) = \pi_2(T^k(x - M_0\omega, \beta)).$$

From (5.2) it follows that $N(x; I_0) \geq M_0$ for all $x \in J_0 - M_0\omega$. Exactly as in the NR case we get

$$(5.10) \quad (x + \omega, y_{M_0+1}(x)) \in S \text{ and } \frac{1}{\kappa} < y'_{M_0+1}(x) < 2\kappa, \quad x \in J_0,$$

as well as the upper bound

$$y'_{M_0+K_0}(x) < \varepsilon^{-\rho K_0} \text{ on } J_0.$$

Since $J_0 \cap I_0^2 = \emptyset$ by (5.2), it follows from Lemma 2.4 that $\{x \in J_0 : y_{M_0+1}(x) \in A'\}$ is a single interval, and $y_{M_0+1}(J_0) \supset A'$.

In this resonant case we will use the shadowing argument twice: one time to get from $J_0 + \omega$ to $(J_0 + (\nu_0 + 1)\omega)$, and a second time to go from $(J_0 + (\nu_0 + 1)\omega)$ to $J_0 + K_0\omega$. First, let, as above, $\eta_k(x) = \pi_2(T^k(x + \omega, \beta))$. Applying Lemma 2.11 we get

$$(5.11) \quad [\eta_0, y_{M_0+1}(s)], [\eta_0, y_{M_0+1}(t)] \subset \mathbb{T} \setminus A,$$

and there is a unique $\xi \in J_0$ such that $y_{M_0+1}(\xi) = \beta = \eta_0$, and y_{M_0+1} (trivially) overtakes η_0 at ξ .

By (5.2) we have $(J_0 + k\omega) \cap I_0 = \emptyset$ for all $1 \leq k \leq \nu_0 - 1$, and $(J_0 + \nu_0\omega) \cap I_0 = I_0^2$. We therefore get, by proceeding exactly as above (with $(x + \omega, \eta_0)$, $x \in J_0$), that

$$(x + (\nu_0 + 1)\omega, \eta_{\nu_0}(x)) \in S, \quad x \in J_0;$$

and $\{x \in J_0 : \eta_{\nu_0}(x) \in A'\}$ is a single interval, and $\eta_{\nu_0}(J_0) \supset A'$. Moreover, since (5.11) holds, we also get

$$\{s + (\nu_0 + 1)\omega\} \times [\eta_{\nu_0}(s), y_{M_0+\nu_0+1}(s)], \{t + (\nu_0 + 1)\omega\} \times [\eta_{\nu_0}(t), y_{M_0+\nu_0+1}(t)] \subset S,$$

and since $s + (\nu_0 + 1)\omega, t + (\nu_0 + 1)\omega \notin \text{int}(I_0 + \omega)$ we have (recall Lemma 2.2)

$$(5.12) \quad [\eta_{\nu_0}(s), y_{M_0+\nu_0+1}(s)], [\eta_{\nu_0}(t), y_{M_0+\nu_0+1}(t)] \subset \overline{\mathbb{T} \setminus A'}.$$

We are thus in a situation where we can apply Lemma 2.12. We conclude that the equation $y_{M_0+\nu_0+1}(x) = \beta$ has exactly two solutions ξ_1, ξ_2 in $J_0 = [s, t]$, both lying in the interior of J_0 , and $y_{M_0+\nu_0+1}(x)$ overtakes the constant function $\sigma_0(x) = \beta$ at these two points. See Fig. 11. Moreover,

$$(5.13) \quad [\beta, y_{M_0+\nu_0+1}(s)], [\beta, y_{M_0+\nu_0+1}(t)] \subset \mathbb{T} \setminus A.$$

We can now use the above control to verify that the set I_1 must be "well-inside" J_0 . We notice that, since (5.2) holds, we get that $N(x, I_0) \gg K_0$ for $x \in J_0 + (\nu_0 + 1)\omega$. Since $(J_0 + K_0\omega) \cap I_0 = \emptyset$, it thus follows (by the same argument as in the previous case) from Lemma 5.1(1) that $I_1 \subset \{x \in J_0 : y_{M_0+\nu_0+1}(x) \in A'\}$. Since we have the derivative estimates in (5.10), it follows from the trivial upper bound in Lemma 3.3(3) that $y'_{M_0+\nu_0+1} < 3\kappa\varepsilon^{-\rho\nu_0}$ on J_0 . Since also (5.12) holds, it follows from Lemma 3.4 that

$$\text{dist}(I_1, \partial J_0) > \varepsilon^{\rho\nu_0+1}/(6\kappa).$$

Finally, to get control on $y_{M_0+K_0}$ we apply shadowing a second time. We shall shadow $y_{M_0+\nu_0+1}$ by the iterates

$$\sigma_k(x) = \pi_2(T^k(x + (\nu_0 + 1)\omega, \beta)).$$

We noted above that $N(x, I_0) \gg K_0$ for $x \in J_0 + (\nu_0 + 1)\omega$. Thus it follows from Lemma 5.1 (since $\beta \in \mathbb{T} \setminus A$) that

$$\sigma_{K_0-(\nu_0+1)}(x) \in \mathbb{T} \setminus A' \text{ for all } x \in J_0;$$

and since (5.13) holds we get

$$[\sigma_{K_0-(\nu_0+1)}(s), y_{M_0+K_0}(s)], [\sigma_{K_0-(\nu_0+1)}(t), y_{M_0+K_0}(t)] \subset \mathbb{T} \setminus A'.$$

Since $y_{M_0+K_0}$ overtakes $\sigma_{K_0-(\nu_0+1)}$ at ξ_1, ξ_2 , we are thus in a situation where we can apply Lemma 2.13(2) and conclude that there are two disjoint intervals I_1^1, I_1^2 in J_0 such that $I_1^1 \cup I_1^2 = \{x \in J_0 : y_{M_0+K_0}(x) \in A'\}$; and $y_{M_0+K_0}(I_1^i) = A'$ ($i = 1, 2$).

What is left now is to get a lower bound on $y'_{M_0+K_0}$ on I_1 . Since $N(x, I_0) \gg K_0$ for all $x \in I_0 + (\nu_0 + 1)\omega$ it follows from Lemma 5.1(2) that if $(\theta, \eta) \in (I_0 + (\nu_0 + 1)\omega) \times \mathbb{T}$ are such that $\eta_{K_0-(\nu_0+1)} \in A'$ then

$$\prod_{j=0}^{K_0-\nu_0-2} f'(\eta_j) > \varepsilon^{-(K_0-\nu_0-1)}.$$

Since we by Lemma 3.3(4) have the trivial bound $y'_{M_0+\nu_0+1}(x) > 1/\kappa$ on \mathbb{T} , it thus follows from Lemma 3.3(2) that $y'_{M_0+K_0}(x) > \varepsilon^{-(K_0-\nu_0-1)}/\kappa \gg \varepsilon^{-K_0/2}$ for a.e. $x \in I_1^1$ (recall that $\nu_0^2 \leq K_0$).

□

Remark 7. As we also remarked earlier, assuming that $g(x)$ having degree > 2 would not add any fundamental differences (we just got longer chains of the shadowing argument when we have, the very unlikely, situations of multiple resonances). However, the notation would become slightly more involved.

6. INDUCTIVE STEP

We are now ready for the the inductive step in the proof of Proposition 4.1. To simplify the statements in the next lemma we shall use the following notation: given sets J_i and positive integers M_i, K_i ($i \geq 0$), we let X_n, X'_n, G_n, H_n ($n \geq 0$) denote the following sets:

$$\begin{aligned} X_n &= \mathbb{T} \setminus \bigcup_{j=0}^n \bigcup_{m=-(M_j-1)}^{K_j-1} (J_j + m\omega), \quad X_{-1} = \mathbb{T}; \\ X'_n &= \mathbb{T} \setminus \bigcup_{j=0}^n \bigcup_{m=-M_j^{3/2}}^{K_j-1} (J_j + m\omega), \quad X'_{-1} = \mathbb{T}; \\ G_n &= \bigcup_{j=0}^n \bigcup_{m=1}^{K_j} (J_j + m\omega), \quad G_{-1} = \emptyset; \quad \text{and} \\ H_n &= \mathbb{T} \setminus \bigcup_{j=1}^n \bigcup_{m=0}^{M_j} (J_j + m\omega), \quad H_{-1} = \mathbb{T}. \end{aligned}$$

We note that $X'_n \subset X_n$.

The (vague) motivation for these sets is that the sets J_i should be thought of as "bad", and we need to control the approach rate to them (under translation by ω), as well as the recovering time after visits.

For consistency in the statement of the next lemma we define

$$M_{-1} = K_{-1} = 1.$$

Moreover, in order to get the notations slightly clearer we shall make use of the following convention: Estimates of products of the form $\prod_{j=k}^\ell f'(y_j)$, should be thought to hold, under appropriate assumptions on (x, y) , whenever the expression exists (which it does for a.e. $y \in \mathbb{T}$).

Lemma 6.1. *There exists an $\varepsilon_2(\gamma, \kappa, \rho) > 0$ such that for all $0 < \varepsilon < \varepsilon_2$ and any $n \geq 0$ we have the following: Assume that the closed sets $J_0 \supset J_1 \supset \dots \supset J_n$ and the integers M_j, K_j ($j = 0, 1, \dots, n$) have been constructed, and that the open sets $\widehat{J}_j \supset J_j$ ($j \in [0, n]$) have been chosen, so that the following four conditions hold:*

Condition (1)_n: For each $j \in [0, n]$ there holds:

$$(6.1) \quad 1/2 < M_j/K_j^2 < 2 \text{ and } [\varepsilon^{-K_{j-1}/160}] \leq K_j \leq 2[\varepsilon^{-K_{j-1}/20}];$$

$$(6.2) \quad \widehat{J}_j \cap (\widehat{J}_j + k\omega) = \emptyset \text{ for } 0 < |k| \leq M_j^2;$$

The sets J_j and \widehat{J}_j both consists of one or two intervals, $J_j^\ell, \widehat{J}_j^\ell$, and the intervals satisfy

$$(6.3) \quad \varepsilon^{2\rho K_{j-1}} < |J_j^\ell| < |\widehat{J}_j^\ell| < \varepsilon^{K_{j-1}/2}$$

and

$$(6.4) \quad 3\widehat{J}_j^\ell \cap (3\widehat{J}_j^\ell + k\omega) = \emptyset \text{ for } 0 < |k| \leq M_j^2.$$

Moreover, if $n \geq 1$, we have

$$(6.5) \quad J_n - M_n\omega, J_n + K_n\omega \subset \mathbb{T} \setminus \bigcup_{j=0}^{n-1} \bigcup_{m=-M_j^{3/2}}^{M_j^{3/2}} (\widehat{J}_j + m\omega) \quad (\subset X'_{n-1}).$$

Condition (2)_n: If $(x, y) \in X_n \times (\mathbb{T} \setminus A)$ and $N = N(x; J_n)$ then:

$$(6.6) \quad \prod_{j=1}^k f'(y_j) < \varepsilon^{(1/2+1/2^{n+1})k} \text{ for all } k \in [1, N];$$

for all $t \in [1, N]$ such that $x_t \in H_{n-1}$ we have

$$(6.7) \quad \prod_{j=k}^t f'(y_j) < \varepsilon^{(1/2+1/2^{n+1})(t-k+1)} \text{ for all } k \in [1, t]; \quad \text{and}$$

$$(6.8) \quad \text{if } y_k \in A' \text{ for some } k \in [1, N], \text{ then } x_k \in G_{n-1}.$$

Condition (3)_n: Assume that $x \in X'_n$. If $(x_k, y_k) \in (\mathbb{T} \setminus G_{n-1}) \times A'$ for some $1 \leq k \leq N(x; J_n)$ and some $y \in \mathbb{T}$, then

$$\prod_{j=0}^{k-1} f'(y_j) > \varepsilon^{-(1/2+1/2^{n+1})k}.$$

Condition $(4)_n$: If $\Gamma_n = (J_n - M_n\omega) \times \{\beta\}$ and

$$I_{n+1} = \pi_1(T^{M_n+K_n}(\Gamma_n)) \cap (\mathbb{T} \times A') - K_n\omega,$$

then $I_{n+1} \subset J_n$ consists of two disjoint intervals I_{n+1}^i ($i = 1, 2$), and $\text{dist}(I_{n+1}, \partial J_n) > \varepsilon^{K_n^{2/3}}$. Moreover, if we write $T^{M_n+K_n}(\Gamma_n) = \{(x + K_n\omega, \varphi_n(x)) : x \in J_n\}$, then $\varphi_n(I_{n+1}^i) = A'$ and $\varepsilon^{-K_n/2} < \varphi_n'(x) < \varepsilon^{-\rho K_n}$ on I_{n+1}^i ($i = 1, 2$).

Then there is a closed set $J_{n+1} \subset J_n$, and an open set $\widehat{J}_{n+1} \supset J_{n+1}$, both consisting of one or two intervals, as well as integers M_{n+1}, K_{n+1} such that $(1-4)_{n+1}$ hold. Furthermore, the following statement holds for each interval J_{n+1}^ℓ in J_{n+1} :

Condition $(C)_n$: If $\mathcal{O} \subset (J_{n+1}^\ell - M_n\omega) \times (\mathbb{T} \setminus A)$ is a set such that $\pi_1(\mathcal{O})$ is δ -dense in $J_{n+1}^\ell - M_n\omega$, then $\pi_2(T^{M_n+K_n}(\mathcal{O}))$ is $(\varepsilon^{-\rho K_n} \cdot \delta + 2\varepsilon^{(M_n+K_n)/2})$ -dense in A' .

Remark 8. We note that properties $(1-3)_n$ are of an arithmetic nature while $(4)_n$ contains information about the geometry. Moreover, condition $(3)_n$, which actually is a statement about iterations of T^{-1} , could have been stated in a more precise way, similar to $(2)_n$. But we do not really need better estimates, so we have chosen to formulate the condition like this. (This is the same comment as the one in the remark following Proposition 4.1.)

Before proving this lemma we will show how it, together with the base case in the previous section, is used to prove Proposition 4.1.

Proof of Proposition 4.1. We recall that $M_{-1} = K_{-1} = 1$. The sets J_0, \widehat{J}_0 and the integers K_0, M_0 were defined in Subsection 5.2. We now check that the premises $(1-4)_n$ in Lemma 6.1 are satisfied for $n = 0$.

From the definitions in Subsection 5.2 we note that condition $(1)_0$ holds (note that condition (6.5) is void in the case $n = 0$). Condition $(2)_0$ easily follows from Lemma 5.1(1). Indeed, if $x \in X_0$ it follows from Lemma 5.2 that $N(x; J_0) \leq N(x; I_0)$. Note that $H_{-1} = G_{-1} = \emptyset$. (In particular, condition $(6.8)_0$ says that if $y \in \mathbb{T} \setminus A$, then $y_k \in A'$ for all $1 \leq k \leq N(x; J_0)$.) Furthermore, condition $(3)_0$ follows from Lemma 5.1(2). Finally the geometric condition $(4)_0$ is the content of Lemma 5.3.

Thus the inductive machine starts and gives us closed sets $J_0 \supset J_1 \supset \dots$, open sets $\widehat{J}_j \supset J_j$ ($j \geq 0$), and integers M_j, K_j ($j \geq 0$) such that conditions $(1-4)_n$ holds for all $n \geq 0$.

We begin by looking at the sizes of the integers M_j, K_j . If we let $\Delta = \varepsilon^{-1/160}$, which can be made as big as we want by taking ε small, we have $K_j \geq \Delta^{K_{j-1}}$ for all $j \in [0, n]$. By (6.1) we see that $M_j \approx K_j^2$. We also note that if we let $N_j = M_j^2$, then

$$(6.9) \quad M_j/N_j < 2/K_j^2 \quad \text{and} \quad M_j^{3/2}/N_j < 2/K_j.$$

Moreover,

$$K_j/M_j < 2/K_j \quad \text{and} \quad M_j/M_j^{3/2} < 2/K_j.$$

We now define $X \subset \mathbb{T}$ to be the set

$$X = \mathbb{T} \setminus \bigcup_{j=0}^{\infty} \bigcup_{m=-M_j^{3/2}}^{K_j} (J_j + m\omega).$$

From $(1)_n$ we see that $M_j < 2K_j^2 < 8\varepsilon^{-K_{j-1}/10}$ and $|J_j| < 2\varepsilon^{K_{j-1}/2}$ for all $j \geq 0$. Thus

$$|X| \geq 1 - \sum_{j=0}^{\infty} 2M_j^{3/2}|J_j| \approx 1.$$

By definition we have $X_n \supset X'_n \supset X$ and $\mathbb{T} \setminus G_n \supset X$ for all $n \geq 0$. We also note that if we fix $x \in X$, then $N(x; J_n) \geq M_n^{3/2}$ for all $n \geq 0$. Therefore statement (1) in Proposition 4.1, and the first half of statement (2), follow immediately from $(2)_n$ ($n \geq 0$) in Lemma 6.1. Moreover, statement (3) in Proposition 4.1 follows from $(3)_n$ ($n \geq 0$) in Lemma 6.1

It remains to prove the second part of (2), i.e., to prove that

$$\overline{\{(x_j, y_j)\}_{j \geq 0}} = \mathbb{T}^2 \text{ for all } (x, y) \in X \times (\mathbb{T} \setminus A).$$

To do this, fix $(x, y) \in X \times (\mathbb{T} \setminus A)$, and let $\mathcal{O} = \{(x_j, y_j)\}_{j \geq 0}$ be the forward orbit of (x, y) . We recall that $A' \supset A$. From $(6.8)_n$ in $(2)_n$ it follows that if $x_j \in A'$ for some $j \geq 1$, then $x_j \in G_\infty = \bigcup_{n \geq 0} G_n$. Moreover, combining $(6.2)_n$ and $(6.5)_n$ gives us

$$(J_n - M_n\omega) \cap G_n = \emptyset.$$

Thus, if $x_j \in J_n - M_n\omega$ and $y_j \in A'$, then $x_j \in G_\infty \setminus G_n = \bigcup_{j=n+1}^{\infty} \bigcup_{m=0}^{K_j} (J_j + m\omega)$. From the estimates in $(1)_n$ it follows that $|G_\infty \setminus G_n| \ll \varepsilon^{K_n/4}$. Moreover, by $(6.3)_n$ each interval in J_n has a length $> \varepsilon^{2\rho K_{n-1}} \gg \varepsilon^{K_n/4}$. We conclude that for each $n \geq 0$ the set $\pi_1(\mathcal{O} \cap ((J_n - M_n\omega) \times (\mathbb{T} \setminus A)))$ is $\varepsilon^{K_n/4}$ -dense in $J_n - M_n\omega$. Hence, for each $n \geq 1$ and each J_n^ℓ it follows from condition $(C)_{n-1}$ that the set

$$(6.10) \quad \pi_2(\mathcal{O} \cap ((J_n^\ell + K_n\omega) \times A')) \text{ is } (\varepsilon^{-\rho K_{n-1}} \cdot \varepsilon^{K_n/4} + 2\varepsilon^{(M_{n-1} + K_{n-1})/2})\text{-dense in } A'.$$

We note that $\varepsilon^{-\rho K_{n-1}} \cdot \varepsilon^{K_n/4} + 2\varepsilon^{(M_{n-1} + K_{n-1})/2} \rightarrow 0$ as $n \rightarrow \infty$.

By compactness of \mathbb{T} there must be a point $x^* \in \mathbb{T}$ and an infinite subsequence n_i such that $\text{dist}(x^*, J_{n_i} + K_{n_i}\omega) \rightarrow 0$ as $i \rightarrow \infty$. From (6.10) it thus follows that

$$(6.11) \quad \overline{\mathcal{O}} \supset \{x^*\} \times A'.$$

We claim that $x^* \in X_n$ for all $n \geq 0$. (This is the only place where we need the open intervals \widehat{J}_n .) Indeed, by $(6.5)_n$ we have that each $J_n + K_n\omega \subset S_n := \mathbb{T} \setminus \bigcup_{j=0}^{n-1} \bigcup_{m=-2M_j}^{2M_j} (\widehat{J}_j + m\omega)$. We note that each S_n is closed, and clearly $S_{n+1} \subset S_n$. Moreover, the estimates in $(1)_n$ show (as we did above for X) that $|\bigcap_{n \geq 0} S_n| \approx 1$. From Lemma 7.4 it therefore follows that $x^* \in \bigcap_{n \geq 1} S_n \subset \bigcap_{n \geq 0} X_n$.

To complete the proof we note that (since $x^* \in \cap_{n \geq 0} X_n$) $N(x^*, J_n) \geq M_n$ for all $n \geq 0$. Thus it follows from $(2)_n$ (applied with $x = x^*$) that

$$\prod_{j=1}^k f'(y_j) < \varepsilon^{k/2} \quad \text{for all } k \geq 1 \text{ and for a.e. } y \in \mathbb{T} \setminus A.$$

Combining this estimate with Lemma 3.1 yields $|\pi_1(T^k(x^*, \mathbb{T} \setminus A))| \rightarrow 0$ as $k \rightarrow \infty$. Thus we must have $|\pi_1(T^k(x^*, A))| \rightarrow 1$ as $k \rightarrow \infty$, and hence also $\overline{\bigcup_{k \geq 0} T^k(x^*, A)} = \mathbb{T}^2$. Since (6.11) holds, we can therefore conclude that $\overline{\mathcal{O}} = \mathbb{T}^2$. \square

Now we turn to the proof of the inductive lemma, Lemma 6.1.

Proof of Lemma 6.1. The proof, which is rather long, is divided into steps. The first thing we do is to make a few observation on the position of translates of the set J_n . From $(6.2)_n$ and $(6.5)_n$, together with the definition of the set X_n , it immediately follows that

$$(6.12) \quad J_n - M_n \omega \subset X_n \text{ and } J_n + K_n \omega \subset X'_n \subset X_n.$$

Furthermore, since $J_{j+1} \subset J_j$ ($j = 0, 1, \dots, n-1$) and since (6.2) holds, we have

$$(6.13) \quad J_n, J_n - M_n \omega \subset H_n \subset \mathbb{T} \setminus G_n.$$

These facts will be used several times below. We also recall the discussion about the sizes of the integers K_j, M_j , and their relationships, from the proof of Proposition 4.1 above.

Step 1 (Construction of J_{n+1} and \widehat{J}_{n+1} , and choice of M_{n+1}, K_{n+1} ; verification of $(1)_{n+1}$): From the estimates in $(4)_n$ it follows that each of the two intervals in I_{n+1} has a length in the interval $(\varepsilon^{2\rho K_n}, \varepsilon^{K_n/2}/2)$. Thus these intervals are much much shorter than the ones in J_j and \widehat{J}_j ($j = 0, \dots, n$), as given by the estimates in $(6.3)_j$.

When we shall choose M_{n+1} and K_{n+1} we will use Lemma 7.3 (with the $Z_j = M_j^{3/2}$). From the above observation on the length of the intervals, together with (6.9) and $(6.4)_j$ ($j = 0, \dots, n$), we see that the premises for Lemma 7.3 are fulfilled for any interval J of length $< \varepsilon^{K_n/2}$. (Here the p in Lemma 7.3 is 2).

We now check whether the two intervals in I_{n+1} in $(4)_n$ are resonant or not. We know that the length of each of the intervals satisfies $< \frac{1}{2}\varepsilon^{K_n/2}$. Thus it follows from Lemma 7.1 that

$$I_{n+1}^\ell \cap (I_{n+1}^\ell + k\omega) = \emptyset \text{ for all } 0 < |k| \leq \sqrt{\gamma} \cdot \varepsilon^{-K_n/4}.$$

Case R (resonant case): If there is a ν_{n+1} , $0 < \nu_{n+1} \leq \varepsilon^{-K_n/40}$, such that

$$I_{n+1}^1 \cap (I_{n+1}^2 - \nu_{n+1}\omega) \neq \emptyset$$

then we let J_{n+1} be the interval $J_{n+1} = I_{n+1}^1 \cup (I_{n+1}^2 - \nu_{n+1}\omega)$ (if the above relation holds with I_{n+1}^1 and I_{n+1}^2 interchanged, we relabel the intervals). Thus the length of J_{n+1} is smaller than $\varepsilon^{-K_n/2}$. Let $K = \lceil \varepsilon^{-K_n/20} \rceil$ and $M = K^2$. Applying Lemma 7.3 with $J = J_{n+1} + K\omega$ shows that there is a $K_{n+1} \in [K+1, K+M_n^2]$ such that $J_{n+1} + K_{n+1}\omega$ satisfy $(6.5)_{n+1}$; and applying Lemma 7.3 with $J = J_{n+1} - M\omega$ gives us an $M_{n+1} \in [M - M_n^2, M - 1]$ such that $J_{n+1} + M_{n+1}\omega$ satisfy $(6.5)_{n+1}$. Since $M_n^2 \approx K_n^4 \ll K$ (by $(6.1)_n$), it follows that

$M_{n+1}/K_{n+1} \approx M/K^2 = 1$. Thus $(6.1)_{n+1}$ holds. Finally, we have $|J_{n+1}| < \varepsilon^{-K_n/2}$ (and J_{n+1} is closed). Let $\widehat{J}_{n+1} \supset J_{n+1}$ be an open interval satisfying $|\widehat{J}_{n+1}| < \varepsilon^{-K_n/2}$ (any such interval works). Since $M_{n+1}^2 \approx \varepsilon^{-K_n/5} \ll \varepsilon^{-K_n/4}$, conditions $(6.2 - 6.4)_{n+1}$ immediately follow from Lemma 7.1. Furthermore, from Lemma 7.2 we get

$$(6.14) \quad \begin{aligned} (J_{n+1} + k\omega) \cap I_{n+1} &= \emptyset \text{ for all } k \in [-M_{n+1}^2, M_{n+1}^2] \setminus \{0, \nu_{n+1}\}, \text{ and} \\ J_{n+1} \cap I_{n+1} &= I_{n+1}^1, (J_{n+1} + \nu_{n+1}\omega) \cap I_{n+1} = I_{n+1}^2. \end{aligned}$$

From the above definitions we see that $\nu_{n+1}^2 \leq K_{n+1}$. Moreover, since $(6.2)_n$ holds, and $I_{n+1} \subset J_n$, we have $\nu_{n+1} \geq M_n^2$. Thus,

$$(6.15) \quad M_n^2 \leq \nu_{n+1} \leq K_{n+1}^{1/2}.$$

Furthermore, since $I_{n+1} \subset J_n$, and since $\text{dist}(I_{n+1}, J_n) > \varepsilon^{K_n^{2/3}} \gg \varepsilon^{K_n/2}$, it follows that

$$(6.16) \quad J_{n+1} = I_{n+1}^1 \cup (I_{n+1}^2 - \nu_{n+1}\omega) \subset J_n \text{ and } J_{n+1} + \nu_{n+1}\omega \subset J_n$$

(since $|I_{n+1}^\ell| < \varepsilon^{K_n/2}$).

Case NR (non-resonant case): Otherwise we have

$$(6.17) \quad I_{n+1} \cap (I_{n+1} + k\omega) = \emptyset \text{ for all } 0 < |k| \leq \varepsilon^{-K_n/40}.$$

In this case we let $J_{n+1} = I_{n+1}$; thus J_{n+1} contains two intervals (and trivially we have $J_{n+1} \subset J_n$). We let $\widehat{J}_{n+1} \supset J_{n+1}$ be an open set consisting of two intervals, each of length $< \varepsilon^{K_n/2}/2$, and such that (6.17) holds with I_{n+1} replaced by \widehat{J}_{n+1} (this is possible since I_{n+1} is closed). Next, let $K = \lceil \varepsilon^{-K_n/160} \rceil$ and $M = K^2$. By proceeding as above, applying Lemma 7.3 with $J = J_{n+1} + K\omega$ and $J = J_{n+1} + M\omega$, respectively, gives us $M_{n+1} \in [M - M_n^2, M - 1]$ and $K_{n+1} \in [K + 1, K + M_n^2]$ such that $(6.5)_{n+1}$ holds. Moreover, since $M_{n+1}^2 \leq \lceil \varepsilon^{-K_n/40} \rceil$, condition $(6.2)_{n+1}$ follows from (6.17); and $(6.4)_{n+1}$ follows from Lemma 7.1 (since $|\widehat{J}_{n+1}^\ell| < \varepsilon^{-K_n/2}/2$).

From these definitions we get, in particular:

Sublemma 1. *If $x \in \mathbb{T} \setminus \bigcup_{m=0}^{K_{n+1}-1} (J_{n+1} + m\omega) \subset X_{n+1}$, then $N(x; J_{n+1}) \leq N(x; I_{n+1})$.*

Proof. Since $\nu_{n+1} \ll K_{n+1}$, the proof is almost identical to that of Lemma 5.2. \square

Step 2 (Verifying $(C)_n$ and $(2)_{n+1}$): First we note that $(1-2)_n$ gives us the following result.

Sublemma 2. *If $x \in J_n - M_n\omega$ then*

$$\prod_{j=1}^k f'(y_j) < \varepsilon^{(1/2+1/2^{n+2})k}$$

for all $k \in [0, M_n + K_n]$ and a.e. $y \in \mathbb{T} \setminus A$.

Proof. From (6.12) we have $x \in J_n - M_n\omega \subset X_n$; thus $N(x; J_n) = M_n$. Hence we can apply $(2)_n$ and conclude that $\prod_{j=1}^k f'(y_j) < \varepsilon^{(1/2+1/2^{n+1})k}$ for all $k \in [0, M_n]$ and a.e. $y \in \mathbb{T} \setminus A$. If $k \in [M_n + 1, M_n + K_n]$ we therefore, trivially, get (since $f'(\eta) < \varepsilon^{-\rho}$ for a.e. $\eta \in \mathbb{T}$) that

$$\prod_{j=1}^k f'(y_j) < \varepsilon^{(1/2+1/2^{n+1})M_n - \rho(k-M_n)}$$

for a.e. $y \in \mathbb{T} \setminus A$. Thus, if we show that $(1/2 + 1/2^{n+1})M_n - \rho(k - M_n) > (1/2 + 1/2^{n+2})k$ we are done.

This inequality can be written $(\rho + 1/2)M_n + M_n/(2^{n+1}) - (\rho + 1/2)k - k/(2^{n+2}) > 0$. Thus the worst case is when $k = M_n + K_n$. For this k the inequality becomes

$$\frac{M_n}{2^{n+2}} - (\rho + 1/2)K_n - \frac{K_n}{2^{n+2}} > 0.$$

Since $\rho + 1/2 + 1/2^{n+2} < 2\rho$, this inequality is fulfilled if $M_n > 2^{n+2} \cdot 2\rho K_n$ (or $K_n/M_n < 1/(2^{n+3}\rho)$). But $K_n/M_n < 2/K_n \ll 1/(2^{n+3}\rho)$, so indeed the statement of the lemma holds. \square

Combining the previous sublemma with Lemma 3.1 immediately gives us

Sublemma 3. *If $x \in J_n - M_n\omega$ and $y, \eta \in \mathbb{T} \setminus A$, then*

$$d(y_{M_n+K_n}, \eta_{M_n+K_n}) \leq \varepsilon^{(1/2+1/2^{n+2})(M_n+K_n)} d(y, \eta).$$

This last sublemma enables us to verify condition $(C)_n$. Assume that $(x, y), (\xi, \eta) \in (J_n^\ell - M_n\omega) \times \mathbb{T} \setminus A$ and that $d(x, \xi) < \delta$. From the upper bound on the derivative of φ_n in condition $(4)_n$ it follows that $d(\pi_2(T^{M_n+K_n}(x, \beta)), \pi_2(T^{M_n+K_n}(\xi, \beta))) < \varepsilon^{-\rho K_n} \cdot \delta$. Moreover, since $\beta \in \mathbb{T} \setminus A$ we can apply Sublemma 3 to get $d(\pi_2(T^{M_n+K_n}(x, \beta)), \pi_2(T^{M_n+K_n}(x, y))) < \varepsilon^{(M_n+K_n)/2}$ and $d(\pi_2(T^{M_n+K_n}(\xi, \beta)), \pi_2(T^{M_n+K_n}(\xi, \eta))) < \varepsilon^{(M_n+K_n)/2}$. We can thus conclude that

$$d(\pi_2(T^{M_n+K_n}(x, y)), \pi_2(T^{M_n+K_n}(\xi, \eta))) < \varepsilon^{-\rho K_n} \cdot \delta + 2\varepsilon^{(M_n+K_n)/2}.$$

This shows that $(C)_n$ holds.

The way to prove $(2)_{n+1}$ is to use the following sublemma. In the proof of this sublemma the "probe" in statement $(4)_n$ will be used.

Sublemma 4. *If $(x, y) \in X_n \times (\mathbb{T} \setminus A)$ and $N = N(x; I_{n+1})$, then the statements (6.6 – 6.8) $_{n+1}$ hold.*

Note that $(2)_{n+1}$ indeed follows from this, since $X_{n+1} \subset X_n$, and since $N(x; J_{n+1}) \leq N(x; I_{n+1})$ for all $x \in X_{n+1}$ by Sublemma 1.

Proof of Sublemma 4. Fix $x \in X_n$ and let $N = N(x; I_{n+1})$. For simplicity we also fix $y \in \mathbb{T} \setminus A$ such that $f'(y_j)$ exists for all $j \geq 0$ (recall that this holds for a.e. y).

The plan is to use $(2)_n$ repeatedly. We let $0 < t_0 < t_1 < \dots < t_s = N$ be the times $\leq N$ when $x_k \in J_n$ (recall that $I_{n+1} \subset J_n$). First we shall use $(2)_n$ to get control up to time t_0 . Then we use the probe from $(4)_n$ to get control on the iterate $(x_{t_0+K_n}, y_{t_0+K_n})$ (we allow

to lose control under the passage from time t_0 to time $t_0 + K_n$). Now we can again (as we will see) use $(2)_n$ (applied to $(x_{t_0+K_n}, y_{t_0+K_n})$) to get information up to time t_1 ; and then we use the probe to get control on $(x_{t_1+K_n}, y_{t_1+K_n})$. Continuing like this we get control up to time $t_s = N$.

Now we show the details in this approach. As above we let $0 < t_0 < t_1 < \dots < t_s = N$ be the times $\leq N$ when $x_k \in J_n$. Note that $t_0 \geq M_n$ since $x \in X_n$, and $t_{j+1} - t_j > M_n^2$ by $(6.2)_n$.

Applying $(2)_n$ to the point (x, y) we get that $(6.6 - 6.8)_n$ hold, with N replaced by t_0 . We denote this by $(6.6 - 6.8)_n[t_0]$. If $t_0 = N$, i.e., if $x_{t_0} \in I_{n+1}$, the (weaker) statements in $(2)_{n+1}$ follow directly from $(2)_n$. If not, we note that $x_{t_0-M_n} \in J_n - M_n\omega$. Since $J_n - M_n\omega \subset \mathbb{T} \setminus G_{n-1}$, by $(6.5)_n$, it follows from $(6.8)_n[t_0]$ that $y_{t_0-M_n} \in \mathbb{T} \setminus A$. Combining $(6.6)_n[t_0]$ with Sublemma 2 gives $(6.6)_{n+1}[t_0 + K_n]$.

Next we recall the definition of Γ_n and the function φ_n in condition $(4)_n$. We let $\eta = \beta$, and we shall compare the iterates of the two points $(x_{t_0-M_n}, y_{t_0-M_n})$ and $(x_{t_0-M_n}, \eta)$, which both are in $(J_n - M_n\omega) \times (\mathbb{T} \setminus A)$. From the definition of I_{n+1} (we have assumed that $x_{t_0} \notin I_{n+1}$) it follows that $\eta_{M_n+K_n} \notin A'$. Moreover, from Sublemma 3 we get that $d(\eta_{M_n+K_n}, y_{t_0+K_n}) \ll \varepsilon$. Recalling the definition of A' in (2.1), we conclude that $y_{t_0+K_n} \in \mathbb{T} \setminus A$, and thus

$$(6.18) \quad (x_{t_0+K_n}, y_{t_0+K_n}) \in (J_n + K_n\omega) \times (\mathbb{T} \setminus A) \subset X_n \times (\mathbb{T} \setminus A),$$

by $(6.12)_n$. We note that $(6.8)_{n+1}[t_0 + K_n]$ trivially follows from $(6.8)_n[t_0]$ since $\bigcup_{m=1}^{K_n} (J_n + m\omega) \subset G_n$. Thus, so far we know that $(6.6)_{n+1}[t_0 + K_n]$ and $(6.8)_{n+1}[t_0 + K_n]$ hold. Moreover, $(6.7)_{n+1}[t_0 + K_n]$ immediately follows from $(6.7)_n[t_0]$, because of the definition of H_n .

Since (6.18) holds, we can proceed exactly as above (taking one more turn, from time $t_0 + K_n$ to time $t_1 > t_0 + M_n^2$, by applying $(2)_n$ to the point $(x_{t_0+K_n}, y_{t_0+K_n})$). We get (from $(6.6 - 6.8)_n$):

$$(6.19) \quad \prod_{j=t_0+K_n+1}^k f'(y_j) < \varepsilon^{(1/2+1/2^{n+1})(k-t_0-K_n)}, k \in [t_0 + K_n + 1, t_1];$$

For all $t \in [t_0 + K_n + 1, t_1]$ such that $x_t \in H_{n-1}$ we have:

$$(6.20) \quad \prod_{j=k}^t f'(y_j) < \varepsilon^{(1/2+1/2^{n+1})(t-k+1)}, k \in [t_0 + K_n + 1, t];$$

and

$$(6.21) \quad \text{if } y_k \in A' \text{ for some } k \in [t_0 + K_n + 1, t_1], \text{ then } x_k \in G_{n-1}.$$

Combining $(6.6)_{n+1}[t_0 + K_n]$ with (6.19) gives $(6.6)_{n+1}[t_1]$; and combining $(6.8)_{n+1}[t_0 + K_n]$ with (6.21) gives $(6.8)_{n+1}[t_1]$.

It just remains to check $(6.7)_{n+1}[t_1]$. Take $t \in [t_0 + K_n + 1, t_1]$ such that $x_t \in H_n \subset H_{n-1}$. Then $t > t_0 + M_n$. We know that (6.20) holds, and from $(6.7)_{n+1}[t_0 + K_n]$ we have, since

(6.13) holds (i.e., $x_{t_0} \in H_n$),

$$\prod_{j=k}^{t_0} f'(y_j) < \varepsilon^{(1/2+1/2^{n+2})(t_0+1-k)}, k \in [0, t_0].$$

Thus, to verify $(6.7)_{n+1}[t_0 + t_1]$ we need to show that

$$\prod_{j=k}^t f'(y_j) < \varepsilon^{(1/2+1/2^{n+2})(t-k+1)}, k \in [t_0 + 1, t_0 + K_n].$$

Take $k \in [t_0 + 1, t_0 + K_n]$. Then, by (6.20), and using that $f'(\eta) < \varepsilon^{-\rho}$ for a.e. $\eta \in \mathbb{T}$, we get

$$\prod_{j=k}^t f'(y_j) = \prod_{j=k}^{t_0+K_n} f'(y_j) \prod_{j=t_0+K_n+1}^t f'(y_j) < \varepsilon^{-\rho(t_0+K_n-k+1)} \cdot \varepsilon^{(1/2+1/2^{n+1})(t-t_0-K_n)}.$$

Recalling that $t > t_0 + M_n$, and by proceeding as in the proof of Sublemma 2 (again using the fact that $K_n^2 \approx M_n$) shows that this last expression is smaller than $\varepsilon^{(1/2+1/2^{n+2})(t-k+1)}$. Thus $(6.7)_{n+1}[t_1]$ holds.

In this way we can continue inductively to show that $(6.6 - 6.8)_{n+1}[t_j]$ hold for $j = 1, 2, \dots, s$. The process stops when we come to time $t_s = N$ (since we then are in I_{n+1} and do not know what happens with the further iterates).

□

Step 3 (Verifying $(3)_{n+1}$): Condition $(3)_{n+1}$ follows directly (by the same argument as the one after the statement of Sublemma 4; recall that $X'_n \subset X_n$) from

Sublemma 5. *Assume that $(x, y) \in X'_n \times \mathbb{T}$ and that for some $1 \leq k \leq N = N(x; I_{n+1})$ we have $(x_k, y_k) \in (\mathbb{T} \setminus G_n) \times A'$. Then*

$$\prod_{j=0}^{k-1} f'(y_j) > \varepsilon^{-(1/2+1/2^{n+2})k}.$$

Proof. Fix $x \in X'_n$, and let $0 < t_0 < t_1 < \dots < t_s = N = N(x; I_{n+1})$ be the times $\leq N$ when $x_k \in J_n$. Be proceeding as in the proof of Sublemma 4, we note that we have $t_0 > M_n^{3/2}$ and $t_{j+1} - t_j > M_n^2$.

Let $(3)_{n+1}[t_j]$ denote the statement in $(3)_{n+1}$ where $N(x; J_{n+1})$ is replaced by t_j . We thus have to show that $(3)_{n+1}[t_s]$ holds. Since $\mathbb{T} \setminus G_n \subset \mathbb{T} \setminus G_{n-1}$ we note that $(3)_{n+1}[t_0]$ follows directly from $(3)_n$ (which we have assumed holds).

Next we shall prove that $(3)_{n+1}[t_1]$ holds. Assume that $(x_k, y_k) \in (\mathbb{T} \setminus G_n) \times A'$ for some $t_0 + 1 \leq k \leq t_1$ and some y . As we noted above we have $t_0 > M_n^{3/2}$. Moreover, since $x_k \notin G_n$, and $k \geq t_0 + 1$ (and $x_{t_0} \in J_n$), we must by the definition of G_n have $k > t_0 + K_n$.

We claim that

$$y_{t_0-M_n} \in A \text{ and } y_{t_0+K_n} \in A.$$

Indeed, if this was not the case we could apply Sublemma 4 (property $(6.8)_{n+1}$) to the points $(x_{t_0-M_n}, y_{t_0-M_n})$ and $(x_{t_0+K_n}, y_{t_0+K_n})$, respectively (recall that (6.12) holds), and conclude that $y_k \notin A'$ (since $x_k \in \mathbb{T} \setminus G_n$).

The product $\prod_{j=0}^{k-1} f'(y_j)$ is now split into three parts:

$$\prod_{j=0}^{k-1} f'(y_j) = \prod_{j=0}^{t_0-M_n-1} \prod_{j=t_0-M_n}^{t_0+K_n-1} \prod_{j=t_0+K_n}^{k-1}$$

Since $y_{t_0-M_n} \in A$ (and since $(J_n - M_n\omega) \cap G_n = \emptyset$), it follows from $(3)_n$, applied to the point (x, y) that $\prod_{j=0}^{t_0-M_n-1} f'(y_j) > \varepsilon^{-(1/2+1/2^{n+1})(t_0-M_n)}$. Moreover, applying $(3)_n$ to the point $(x_{t_0+K_n}, y_{t_0+K_n})$ (recall that (6.12) holds) gives $\prod_{j=t_0+K_n}^{k-1} f'(y_j) > \varepsilon^{-(1/2+1/2^{n+1})(k-(t_0+K_n))}$. Since we have the global bound $f'(\eta) > \varepsilon^\rho$, we thus get

$$\prod_{j=0}^{k-1} f'(y_j) > \varepsilon^{-(1/2+1/2^{n+1})(k-(M_n+K_n))+\rho(M_n+K_n)}$$

Now we use the fact that $k > t_0 + K_n > M_n^{3/2} + K_n > M_n^{3/2}$ to conclude (by using the same type of argument as in the proof of Sublemma 3) that right hand side of this last expression is larger than $\varepsilon^{-(1/2+1/2^{n+2})k}$. We have thus shown that $(3)_{n+1}[t_1]$ holds.

We continue by showing that $(3)_{n+1}[t_2]$ holds. Assume that $(x_k, y_k) \in (\mathbb{T} \setminus G_n) \times A'$ for some $t_1 + 1 \leq k \leq t_2$ and some y . We note that we must have $k > t_1 + K_n$. By the same argument as above we can conclude that $y_{t_i+K_n}, y_{t_i-M_n} \in A$ ($i = 1, 2$), for otherwise we cannot have $y_k \in A'$. We now split the product $\prod_{j=0}^{k-1} f'(y_j)$ into five parts: $\prod_{j=0}^{t_0-M_n-1} \prod_{j=t_0-M_n}^{t_0+K_n-1} \prod_{j=t_0+K_n}^{t_1-M_n-1} \prod_{j=t_1-M_n}^{t_1+K_n-1} \prod_{j=t_1+K_n}^k$. Applying $(3)_n$ to the points $(x, y), (x_{t_i+K_n}, y_{t_i+K_n})$ ($i=1,2$), and using the global lower bound on f' , we get (as above) the estimate

$$\prod_{j=0}^{k-1} f'(y_j) > \varepsilon^{-(1/2+1/2^{n+1})(k-2(M_n+K_n))+2\rho(M_n+K_n)} > \varepsilon^{-(1/2+1/2^{n+2})k}$$

since $k > t_1 + K_n > t_0 + M_n^2 + K_n \gg 2M_n^{3/2}$. Thus $(3)_{n+1}[t_2]$ also holds.

Proceeding inductively (considering the time intervals $[t_2, t_3], \dots, [t_{s-1}, t_s]$) shows that $(3)_{n+1}[t_s]$ holds. \square

Step 4 (Verifying $(4)_{n+1}$). Now we shall control the geometry of the iterates of the segment $\Gamma_{n+1} = (J_{n+1} - M_{n+1}\omega) \times \{\beta\}$. As in the base case we get two cases, depending on whether the intervals I_{n+1} were resonant or not (the analysis in step 1 above). We begin with the resonant case (R) where $J_{n+1} = I_{n+1}^1 \cup (I_{n+1}^2 - \nu_{n+1}\omega)$ is a single interval. This is the more involved case since we need to use "shadowing" twice.

Case R: To verify $(4)_{n+1}$ we need to control $T^{M_{n+1}+K_{n+1}}(\Gamma_{n+1})$. We will take the following route: control $T^{M_{n+1}-M_n}(\Gamma_{n+1})$, use $(4)_n$ to control $T^{M_{n+1}+K_n}(\Gamma_{n+1})$, control $T^{M_{n+1}+\nu_{n+1}-K_n}(\Gamma_{n+1})$, again use $(4)_n$ to control $T^{M_{n+1}+\nu_{n+1}+M_n}(\Gamma_{n+1})$, and finally control $T^{M_{n+1}+K_{n+1}}(\Gamma_{n+1})$.

For convenience we shall collect some of the set inclusions which will be used along the argument. First we recall (6.16), that $J_{n+1}, J_{n+1} + \nu_{n+1}\omega \subset J_n$. Thus, by (6.12) we have

$$J_{n+1} + K_n\omega, J_{n+1} + (\nu_{n+1} + K_n)\omega \subset X'_n.$$

Furthermore, by (6.13) (recall the definitions of G_n and H_n) we have

$$J_{n+1} \subset H_n \text{ and } J_{n+1} - M_n\omega, J_{n+1} + (\nu_{n+1} - M_n)\omega \subset \mathbb{T} \setminus G_n.$$

We also recall that (6.5) $_{n+1}$ holds, and conclude that

$$J_{n+1} - M_{n+1}\omega \subset X_n, \quad J_{n+1} + K_{n+1}\omega \subset H_n \subset \mathbb{T} \setminus G_n.$$

Finally, since we are in the resonant case (R) we note that, by (6.14),

$$(6.22) \quad J_{n+1} \cap I_{n+1} = I_{n+1}^1, \quad (J_{n+1} + \nu_{n+1})\omega \cap I_{n+1} = I_{n+1}^2.$$

We now turn to the analysis. We recall Lemma 2.5 and Lemma 3.3(4), i.e., T preserves orientation and increasing functions are mapped to strictly increasing functions. Moreover, we also recall that if ψ_0 is a strictly increasing function which intersects the constant function $\eta_0 = \beta$ at $x = \xi$, so trivially ψ overtakes η_0 at $x = \xi$, then, by Lemma 2.10, the same holds for the iterates, i.e., ψ_k overtakes η_k at $x = \xi$ for all k . These facts will be used several times below, without explicitly writing it out.

Let

$$y_k(x) = \pi_2(T^k(x - M_{n+1}\omega, \beta)).$$

Then we can write

$$T^k(\Gamma_{n+1}) = \{(x + (k - M_{n+1})\omega, y_k(x)) : x \in J_{n+1}\}$$

In particular we have $T^{M_{n+1}+K_{n+1}}(\Gamma_{n+1}) = \{(x + K_{n+1}\omega, y_{M_{n+1}+K_{n+1}}(x)) : x \in J_{n+1}\}$. Thus, we need to control $y_{M_{n+1}+K_{n+1}}$ for $x \in J_{n+1}$. We shall write $J_{n+1} = [s, t]$.

We first derive an upper bound on $y'_{M_{n+1}+K_{n+1}}(x)$. Since $y_0 = \beta \notin A$ and $J_{n+1} - M_{n+1}\omega \subset X_n$ we can apply Sublemma 4 and derive (since $J_{n+1} \subset H_n$ and $N(x; I_{n+1}) \geq M_{n+1}$ for all $x \in J_{n+1} - M_{n+1}\omega$)

$$\prod_{j=k}^{M_{n+1}} f'(y_j(x)) < \varepsilon^{(M_{n+1}-k+1)/2} \text{ on } J_{n+1}, 1 \leq k \leq M_{n+1}.$$

Below we shall use the fact that Sublemma 4 also gives, since $J_{n+1} - M_n\omega \subset \mathbb{T} \setminus G_n$,

$$(6.23) \quad y_{M_{n+1}-M_n}(x) \in \mathbb{T} \setminus A' \text{ for all } x \in J_{n+1}.$$

Applying Lemma 3.3(1) gives us $y'_{M_{n+1}+1}(x) < 2\kappa$ on J_{n+1} . Hence (trivially)

$$(6.24) \quad y'_{M_{n+1}+k}(x) < 3\kappa\varepsilon^{-\rho(k-1)} < \varepsilon^{-\rho k} \text{ on } J_{n+1} \text{ for any } k \geq 1$$

by Lemma 3.3(3).

From (6.23) we have some control on $y_{M_{n+1}-M_n}$ on J_{n+1} . Next we want to control $y_{M_{n+1}+K_n}$. To do this we use shadowing, i.e., we shall compare the iterates of the points $(x - M_n\omega, y_{M_{n+1}-M_n}(x))$ by those of $(x - M_n\omega, \beta)$, for $x \in J_{n+1}$. Let

$$\eta_k(x) = \pi_2(T^k(x - M_n\omega, \beta)).$$

From $(4)_n$ we have $\eta_{M_n+K_n}(I_{n+1}^1) = A'$ and, since (6.22) holds, we have $\eta_{M_n+K_n}(x) \notin A'$ for all $x \in J_{n+1} \setminus I_{n+1}^1$. Applying Sublemma 3 (recall that (6.23) holds) give us that $y_{M_{n+1}+K_n} \approx \eta_{M_n+K_n}$ on J_{n+1} . Thus we can use Lemma 2.11 and conclude that there is a unique point ξ in $J_{n+1} = [s, t]$, lying in the interior of J_{n+1} , such that

$$(6.25) \quad y_{M_{n+1}+K_n}(\xi) = \beta,$$

and

$$(6.26) \quad [\beta, y_{M_{n+1}+K_n}(s)], [\beta, y_{M_{n+1}+K_n}(t)] \subset \mathbb{T} \setminus A$$

(recall the definition of $A'' \supset A$ in (2.3)).

Now we shall get a control on $y_{M_{n+1}+(\nu_{n+1}-M_n)}$. Recall from (6.15) that we have $\nu_{n+1} \geq M_n^2$. Let

$$\eta_k(x) = \pi_2(T^k(x + K_n\omega, \beta))$$

(we "recycle" the variables η_k in order not to have too many variables). Since $J_{n+1} + K_n\omega \subset X_n$ and $J_{n+1} + (\nu_{n+1} - M_n)\omega \subset \mathbb{T} \setminus G_n$, and since $N(x; I_{n+1}) \geq (\nu_{n+1} - K_n)$ for all $x \in J_{n+1} + K_n\omega$, it follows from Sublemma 4 that $\eta_{\nu_{n+1}-K_n-M_n}(x) \in \mathbb{T} \setminus A'$ for all $x \in J_{n+1}$. Moreover, since (6.26) holds the same argument also shows that

$$[\eta_{\nu_{n+1}-K_n-M_n}(x), y_{M_{n+1}+\nu_{n+1}-M_n}(x)] \subset \mathbb{T} \setminus A' \text{ for } x = s, t.$$

This last inclusion trivially implies (see Lemma 2.1) that

$$(6.27) \quad [\beta, y_{M_{n+1}+\nu_{n+1}-M_n}(s)], [\beta, y_{M_{n+1}+\nu_{n+1}-M_n}(t)] \subset \mathbb{T} \setminus A.$$

Since (6.25) holds we also note that $y_{M_{n+1}+\nu_{n+1}-M_n}(x)$ overtakes $\eta_{\nu_{n+1}-K_n-M_n}(x)$ at $x = \xi$ (recall Lemma 2.10). Applying Lemma 2.13(1), with $I = J_{n+1}$, shows that there is a unique point $\tilde{\xi} \in J_{n+1}$ such that $y_{M_{n+1}+\nu_{n+1}-M_n}(\tilde{\xi}) = \beta$.

To estimate $y_{M_{n+1}+\nu_{n+1}+K_n}$ we use shadowing again. More precisely, we shall compare the iterates of $(x + (\nu_{n+1} - M_n)\omega, y_{M_{n+1}+\nu_{n+1}-M_n}(x))$ by the iterates of $(x + (\nu_{n+1} - M_n)\omega, \beta)$. Therefore we let (again recycling the variables η_k)

$$\eta_k(x) = \pi_2(T^k(x + (\nu_{n+1} - M_n)\omega, \beta)).$$

We recall that $J_{n+1} + (\nu_{n+1} - M_n)\omega \subset J_n - M_n\omega$ and $(J_{n+1} + \nu_{n+1}\omega) \cap I_{n+1} = I_{n+1}^2$. Thus it follows from $(4)_n$ that $\eta_{M_n+K_n}(I_{n+1}^2 - \nu_{n+1}\omega) = A'$ and we have $\eta_{M_n+K_n}(x) \notin A'$ for all $x \in J_{n+1} \setminus (I_{n+1}^2 - \nu_{n+1}\omega)$. In particular we have $\eta_{M_n+K_n}(x) \notin \text{int}(A')$ for $x = s, t$. We also note that, by the definition of $\tilde{\xi}$ above, $y_{M_{n+1}+\nu_{n+1}+K_n}$ overtakes $\eta_{M_n+K_n}$ at $x = \tilde{\xi}$. Moreover, using (6.27) and Sublemma 3 we conclude that

$$[\eta_{M_n+K_n}(x), y_{M_{n+1}+\nu_{n+1}+K_n}(x)] \subset \mathbb{T} \setminus A'' \text{ for } x = s, t.$$

We are thus in a position where we can apply Lemma 2.12 and conclude that there are exactly two $\xi_1, \xi_2 \in J_{n+1}$ such that $y_{M_{n+1}+\nu_{n+1}+K_n}(\xi_i) = \beta$, both lying in the interior of J_{n+1} , and

$$(6.28) \quad [\beta, y_{M_{n+1}+\nu_{n+1}+K_n}(x)] \subset \mathbb{T} \setminus A \text{ for } x = s, t.$$

To show that I_{n+2} , as defined in $(4)_{n+1}$, is "well inside" J_{n+1} we shall also use the following easy observation. From (6.24) we know that $y'_{M_{n+1}+\nu_{n+1}+K_n}(x) < \varepsilon^{-\rho(\nu_{n+1}+K_n)}$ on J_{n+1} . It thus follows from Lemma 3.4 that

$$\text{dist}(\tilde{I}_{n+2}, \partial J_{n+1}) > \varepsilon^{\rho(\nu_{n+1}+K_n)+1}/2 \gg \varepsilon^{K_{n+1}^{2/3}}$$

where $\tilde{I}_{n+2} = \{x \in J_{n+1} : y_{M_{n+1}+\nu_{n+1}+K_n}(x) \in A\}$. In the last inequality we used (6.15).

The last step is to control $y_{M_{n+1}+K_{n+1}}$. Let

$$\eta_k(x) = \pi_2(T^k(x + (\nu_{n+1} + K_n)\omega, \beta)).$$

As we recalled above we have $J_{n+1} + (\nu_{n+1} + K_n)\omega \subset X_n$ and $J_{n+1} + K_{n+1}\omega \subset \mathbb{T} \setminus G_n$, and by (6.14) we have $N(x; I_{n+1}) \gg K_{n+1}$ for all $x \in J_{n+1} + (\nu_{n+1} + K_n)\omega$. Therefore it follows from Sublemma 4 that

$$\eta_{K_{n+1}-\nu_{n+1}-K_n}(x) \notin A' \text{ for all } x \in J_{n+1};$$

and by the definitions of \tilde{I}_{n+2} above we get that $y_{M_{n+1}+K_{n+1}}(x) \notin A'$ for all $x \in J_{n+1} \setminus \tilde{I}_{n+2}$. Hence

$$I_{n+2} \subset \tilde{I}_{n+2}.$$

We also get, since (6.28) holds,

$$[\eta_{K_{n+1}-\nu_{n+1}-K_n}(x), y_{M_{n+1}+K_{n+1}}(x)] \subset \mathbb{T} \setminus A' \text{ for } x = s, t.$$

We can therefore apply Lemma 2.13(2) (recall that $y_{M_{n+1}+K_{n+1}}$ overtakes $\eta_{K_{n+1}-\nu_{n+1}-K_n}$ at $x = \xi_1, \xi_2$) and conclude that $\{x \in J_{n+1} : y_{M_{n+1}+K_{n+1}}(x) \in A'\}$ consists of two intervals, $I_{n+2}^i (i = 1, 2)$, and $y_{M_{n+1}+K_{n+1}}(I_{n+2}^i) = A' (i = 1, 2)$.

It remains to get a lower bound on the derivative of $y_{M_{n+1}+K_{n+1}}$ on I_{n+2} . To obtain this we use Sublemma 5 for the passage from $J_{n+1} + (\nu_{n+1} + K_n)\omega$ to $J_{n+1} + K_{n+1}\omega$. We conclude that if $y_{M_{n+1}+K_{n+1}}(x) \in A'$ for some $x \in J_{n+1} + (\nu_{n+1} + K_n)\omega \subset J_n + K_n\omega \subset X'_n$, then

$$\prod_{k=\nu_{n+1}+K_n}^{K_{n+1}-1} f'(y_{M_{n+1}+k}(x)) > \varepsilon^{-(1/2+1/2^{n+2})(K_{n+1}-(\nu_{n+1}+K_n))} \gg \kappa \varepsilon^{-K_{n+1}/2}$$

whenever the derivatives exist. The last inequality follows from the fact that (6.15) holds, i.e., $\nu_{n+1}^2 \leq K_{n+1}$, and we have $K_{n+1}^{1/2}/2^{n+2} \gg 1$ (and $K_n \ll K_{n+1}$). By applying Lemma 3.3(2), recalling the (global) trivial lower bound $y'_{M_{n+1}+\nu_{n+1}+K_n}(x) > 1/\kappa$ given by Lemma 3.3(4), we obtain the desired lower bound on $y'_{M_{n+1}+K_{n+1}}$ on I_{n+2} .

Case NR: In this case $J_{n+1} = I_{n+1}$ consist of two intervals, and we only need to use shadowing once (since there is no interactions). We can treat these two intervals independently, like we did in the proof of Lemma 5.3. We focus on one of them, say $I_{n+1}^1 = [s, t]$. The route we shall take is the following: control $T^{M_{n+1}-M_n}(\Gamma_{n+1}^1)$, use $(4)_n$ to control $T^{M_{n+1}+K_n}(\Gamma_{n+1}^1)$, and control $T^{M_{n+1}+K_{n+1}}(\Gamma_{n+1}^1)$.

As before we let

$$y_k(x) = \pi_2(T^k(x - M_{n+1}\omega, \beta)),$$

and we need estimates of $y_{M_{n+1}+K_{n+1}}$. By proceeding exactly as above we get the following estimates on $y_{M_{n+1}+K_n}$: there is a unique point $\xi \in I_{n+1}^1 = [s, t]$, lying in the interior of I_{n+1}^1 , such that $y_{M_{n+1}+K_n}(\xi) = \beta$, and $[\beta, y_{M_{n+1}+K_n}(s)], [\beta, y_{M_{n+1}+K_n}(t)] \subset \mathbb{T} \setminus A$. We let

$$\eta_k(x) = \pi_2(T^k(x + K_n\omega, \beta)).$$

Since $I_{n+1}^1 + K_n\omega \subset X_n$ and $N(x; I_{n+1}) \gg K_{n+1} - K_n$ for all $x \in I_{n+1} + K_n\omega$ we can apply Sublemma 4 to conclude that $\eta_{K_{n+1}-K_n}(x) \notin A'$ for all $x \in I_{n+1}^1$. By applying Lemma 2.13(1) we see that $\{x \in I_{n+1}^1 : y_{M_{n+1}+K_{n+1}}(x) \in A'\}$ consists of one interval, I_{n+2}^1 , and $y_{M_{n+1}+K_{n+1}}(I_{n+2}^1) = A'$. It also follows that $I_{n+2}^1 \subset \{x \in I_{n+1}^1 : y_{M_{n+1}+K_n}(x) \in A\}$ which is "well" inside I_{n+1}^1 , as in the previous case. Proceeding as in the resonant case above we obtain the desired lower bound on the derivative of $y_{M_{n+1}+K_{n+1}}$ on I_{n+2}^1 .

This finishes the proof of the inductive lemma. \square

7. MISCELLANEOUS

In this section we have collected results of a computational nature. The first three lemmas are related to number theory; the last one is from ergodic theory.

The first lemma gives a lower bound on the return-time to an interval under translation by a Diophantine number.

Lemma 7.1. *Assume that ω satisfies the Diophantine condition (1.3) for some $\gamma > 0$. If $I \subset \mathbb{T}$ is an interval of length $\leq \delta$, then*

$$I \cap (I + k\omega) = \emptyset$$

for all $0 < |k| \leq \sqrt{\gamma/\delta}$.

Proof. If $I \cap (I + k\omega) \neq \emptyset$ we must have $|k\omega - p| \leq \delta$ for some integer p . Thus, using (1.3) we get that either $\gamma/k^2 < \delta$, i.e., $|k| > \sqrt{\gamma/\delta}$, or $k = 0$. \square

Lemma 7.2. *Assume that ω satisfies the Diophantine condition (1.3) for some $\gamma > 0$, and assume that $I^1, I^2 \subset \mathbb{T}$ are two disjoint intervals of length $\leq \delta$. Let $N = \lceil \sqrt{\gamma/(2\delta)} \rceil$ and assume that there is an integer ν , $0 < \nu < N$, such that $I^1 \cap (I^2 - \nu\omega) \neq \emptyset$. If we let J be the interval $J = I^1 \cup (I^2 - \nu\omega)$, then*

$$(J + k\omega) \cap I = \emptyset \text{ for all } k \in [\nu - N, N] \setminus \{0, \nu\},$$

where $I = I^1 \cup I^2$. Moreover, $J \cap I = I^1$ and $(J + \nu\omega) \cap I = I^2$.

Proof. Since $|J| \leq 2\delta$, it follows from Lemma 7.1, and the choice of N , that $(J + k\omega) \cap J = \emptyset$ for all $0 < |k| \leq N$. Moreover, $(J + k\omega) \cap (J + \nu\omega) \neq \emptyset$ iff $(J + (k - \nu)\omega) \cap J = \emptyset$. Thus, by the same argument we have $(J + k\omega) \cap (J + \nu\omega) = \emptyset$ for all $0 < |k - \nu| \leq N$. From this it now follows (since $I^1 \subset J$ and $I^2 \subset J + \nu\omega$) that $(J + k\omega) \cap I = \emptyset$ for all $k \in [\nu - N, N] \setminus \{0, \nu\}$ (recall that $0 < \nu < N$). This shows the first part of the lemma-

To finish we first note that (again by Lemma 7.1) $(I^i + k\omega) \cap I^i = \emptyset$ for all $0 < |k| \leq N$ ($i = 1, 2$). Thus, we must have $J \cap I = (I_1 \cup (I_2 - \nu\omega)) \cap (I_1 \cup I_2) = I_1$ and $(J + \nu\omega) \cap I = ((I_1 + \nu\omega) \cup I_2) \cap (I_1 \cup I_2) = I_2$. \square

The following lemma will be used in the inductive step in the proof of Proposition 4.1 to show that there is always "space" for a small set consisting of tiny intervals to be translated into a "good" position.

Lemma 7.3. *Assume that each $I_j \subset \mathbb{T}$ ($j = 0, 1, \dots, n$) consists of at most $p \geq 1$ intervals. Assume further that there are integers $N_0 < N_1 < \dots < N_n$ such that for each interval I_j^ℓ in I_j ($j \in [0, n]$) we have*

$$(7.1) \quad 3I_j^\ell \cap \bigcup_{m=1}^{N_j} (3I_j^\ell + m\omega) = \emptyset.$$

If the set $J \subset \mathbb{T}$ consists of at most p intervals, $\max_i |J^i| < \min_{j,\ell} |I_j^\ell|$, and if the integers $Z_j > 0$ are chosen so that $10p^2 Z_j / N_j < 3^{-(j+1)}$, then there is an integer $k \in [1, N_n]$ (in fact at least $N_n/2$ of them) such that

$$(7.2) \quad (J + k\omega) \subset \mathbb{T} \setminus \bigcup_{j=0}^n \bigcup_{m=-Z_j}^{Z_j} (I_j + m\omega).$$

Proof. We first fix $j \in [0, n]$ and pick one interval J^i from J , and one from I_j , denoted I_j^ℓ . From condition (7.1), and the fact that $|J^i| < |I_j^\ell|$ (so if $(J^i + k\omega) \cap I_j^\ell \neq \emptyset$, then $(J^i + k\omega) \subset 3I_j^\ell$), we deduce that in the interval $[1, N_j]$ there are at most $2Z_j + 1$ integers k such that

$$(J^i + k\omega) \cap \bigcup_{m=-Z_j}^{Z_j} (I_j^\ell + m\omega) \neq \emptyset.$$

Therefore, in $[1, N_n]$ there are at most $([N_n/N_j] + 1)(2Z_j + 1) < N_n(10Z_j/N_j)$ integers k such that the above condition holds. Since J and I_j each consists of at most p intervals, there are at most $p^2 N_n(10Z_j/N_j)$ integers $k \in [1, N_n]$ such that

$$(J + k\omega) \cap \bigcup_{m=-Z_j}^{Z_j} (I_j + m\omega) \neq \emptyset.$$

Consequently, using the estimates on $10Z_j/N_j$ we conclude that there are at most

$$p^2 N_n (10Z_0/N_0 + 10Z_1/N_1 + \dots + 10Z_n/N_n) < N_n \sum_{j=0}^n \frac{1}{3^{j+1}} < N_n/2$$

integers k in $[1, N_n]$ for which (7.2) is violated. Thus, the statement of the lemma holds. \square

The next lemma is an elementary topological fact (which we use in the proof of Proposition 4.1).

Lemma 7.4. *Assume that $\emptyset \neq J_n \subset S_n \subset \mathbb{T}$ ($n \geq 0$) are closed sets such that $S_{n+1} \subset S_n$ for all $n \geq 0$. If $x^* \in \mathbb{T}$ is such that $\text{dist}(x^*, J_n) \rightarrow 0$ as $n \rightarrow \infty$, then $x^* \in \bigcap_{n \geq 0} S_n$.*

Proof. We prove this by contradiction. Assume that $x^* \notin \bigcap_{n \geq 0} S_n$. Then there is an $m \geq 0$ such that $x^* \notin S_m$, and thus $x^* \notin S_n$ for all $n \geq m$. Since S_m is closed we must have $\text{dist}(x^*, S_m) > \delta$ for some $\delta > 0$. But this implies $\text{dist}(x^*, S_n) > \delta$ for all $n \geq m$, and hence $\text{dist}(x^*, J_n) > \delta$ for all $n \geq m$. \square

The last proposition is due to Herman [13, 14] and Johnson [15]. However, the formulation is slightly different, so for completeness we include a proof.

Proposition 7.5. *Assume that $\omega \in \mathbb{T}$ is irrational and that the map $F : \mathbb{T}^2 \rightarrow \mathbb{T}^2$ is of the form $F(x, y) = (x + \omega, h(x, y))$, where $h : \mathbb{T}^2 \rightarrow \mathbb{T}$ is continuous. Assume further that there are two measurable functions $w^\pm : \mathbb{T} \rightarrow \mathbb{T}$ such that $w^+(x) \neq w^-(x)$ for a.e. $x \in \mathbb{T}$,*

$$(7.3) \quad F(x, w^\pm(x)) = (x + \omega, w^\pm(x + \omega)) \text{ a.e. } x \in \mathbb{T},$$

and such that for a.e. $x \in \mathbb{T}$ there holds

$$(7.4) \quad d(y_n, w^\pm(x_n)) \rightarrow 0 \text{ as } n \rightarrow \infty \text{ for all } y \neq w^\pm(x).$$

Then F has exactly two invariant ergodic Borel probability measures μ^\pm ; and μ^\pm is the push-forward of the Lebesgue measure on \mathbb{T} by the map $x \mapsto (x, w^\pm(x))$. Moreover, F has either only one or only two minimal sets; either $M = \text{supp } \mu^+ = \text{supp } \mu^-$, or $M^\pm = \text{supp } \mu^\pm$.

Proof. Let μ^\pm be the push-forward of the Lebesgue measure on \mathbb{T} (which is unique probability measure on \mathbb{T} invariant under translation $x \mapsto x + \omega$) by the maps $x \mapsto (x, w^\pm(x))$. From (7.3) it easily follows that μ^\pm are F -invariant ergodic Borel probability measures on \mathbb{T}^2 .

We shall now show that these are the only such measures. First we note that by Birkhoff's ergodic theorem the following holds for all $\varphi \in C^0(\mathbb{T}^2)$:

$$(7.5) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(F^k(x, w^\pm(x))) = \int_{\mathbb{T}^2} \varphi d\mu^\pm, \text{ for a.e. } x \in \mathbb{T}.$$

Using (7.4) we note that for a.e. $x \in \mathbb{T}$ and all $y \neq w^-(x)$ we have

$$(7.6) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \left(\sum_{k=0}^{n-1} \varphi(F^k(x, w^+(x))) - \sum_{k=0}^{n-1} \varphi(F^k(x, y)) \right) = 0$$

for all $\varphi \in C^0(\mathbb{T}^2)$. Thus, for a.e. $x \in \mathbb{T}$ we have

$$(7.7) \quad \sum_{k=0}^{n-1} \varphi(F^k(x, y)) = \begin{cases} \int_{\mathbb{T}^2} \varphi d\mu^+, & \text{if } y \neq w^-(x) \\ \int_{\mathbb{T}^2} \varphi d\mu^-, & \text{if } y = w^-(x). \end{cases}$$

Assume that μ is an F -invariant ergodic Borel probability measure. Assume also that $\mu \neq \mu^-$. We shall show that then $\mu = \mu^+$. Since μ and μ^- are both ergodic they must be mutually singular. Thus we can find a set $A \subset \mathbb{T}^2$ of full μ -measure such that $\mu^-(A) = 0$. The projection of μ onto the base, i.e., $\lambda(I) = \mu(\pi_1^{-1}(I))$, is a probability measure on \mathbb{T} which is invariant under $x \mapsto x + \omega$. Hence it must be the Lebesgue measure. It thus

follows that $|\pi_1(A)| = 1$. Since $\mu^-(A) = 0$ we must have $(x, w^-(x)) \notin A$ for a.e. $x \in \mathbb{T}$. It therefor follows from Birkhoff's ergodic theorem and (7.7) that we for all $\varphi \in C^0(\mathbb{T}^2)$ have

$$\int_{\mathbb{T}^2} \varphi d\mu = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(F^k(x, y)) = \int_{\mathbb{T}^2} \varphi d\mu^+ \text{ for } \mu\text{-a.e. } (x, y) \in A.$$

We thus conclude that $\int_{\mathbb{T}^2} \varphi d\mu = \int_{\mathbb{T}^2} \varphi d\mu^+$ for all $\varphi \in C^0(\mathbb{T}^2)$. Therefore, by Riesz representation theorem, we have $\mu = \mu^+$.

We now turn to the minimal sets. Firstly, since F is continuous and \mathbb{T}^2 is compact, we know that F has a minimal set M (in particular we have $F(M) = M$). Consider the restriction $F|_M : M \rightarrow M$. Since M is compact, this map has an $F|_M$ -invariant ergodic Borel probability measure $\tilde{\mu}$ on M . Extending this measure to \mathbb{T}^2 (by letting $\mu(A) = \tilde{\mu}(A \cap M)$) gives us an F -invariant ergodic Borel probability measure μ . Hence we must have $\mu = \mu^+$ or $\mu = \mu^-$. Since the measure μ is supported on M , and since the support is a closed and F -invariant set, and since M is minimal, we must have $\text{supp } \mu = M$.

From this we conclude that F can have at most two minimal sets, since each minimal set supports an F -invariant ergodic Borel probability measure. We also note that the support of an F -invariant measure is closed and F -invariant; hence it contains a minimal set. Thus, either F has a unique minimal set M , and $M = \text{supp } \mu^+ = \text{supp } \mu^-$, or F has two minimal sets, M^\pm , and $M^\pm = \text{supp } \mu^\pm$.

□

REFERENCES

- [1] Avila, A.; Global theory of one-frequency Schrödinger operators. *Acta Math.* 215 (2015), no. 1, 1–54.
- [2] Avila, A.; Krikorian, R.: Monotonic cocycles. *Invent. Math.* 202 (2015), no. 1, 271–331.
- [3] Benedicks, M.; Carleson, L.: The dynamics of the Hénon map. *Ann. of Math.* (2) 133 (1991), no. 1, 73–169.
- [4] Bjerklöv, K.: The dynamics of a class of quasi-periodic Schrödinger cocycles. *Ann. Henri Poincaré* 16 (2015), no. 4, 961–1031.
- [5] Bjerklöv, K.: Positive Lyapunov exponent and minimality for a class of one-dimensional quasi-periodic Schrödinger equations. *Ergodic Theory Dynam. Systems* 25 (2005), no. 4, 1015–1045.
- [6] Bjerklöv, K.: Quasi-periodic perturbation of unimodal maps exhibiting an attracting 3-cycle. *Nonlinearity* 25 (2012), no. 3, 683–741.
- [7] Bourgain, J.; Goldstein, M.; Schlag, W.: Anderson localization for Schrödinger operators on \mathbb{Z} with potentials given by the skew-shift. *Comm. Math. Phys.* 220 (2001), no. 3, 583–621.
- [8] Bowen, R.: Entropy for group endomorphisms and homogeneous spaces. *Trans. Amer. Math. Soc.* 153 1971 401–414.
- [9] Brin, M.; Stuck, G.: *Introduction to Dynamical Systems*. Cambridge University Press, Cambridge, 2002. xii+240 pp. ISBN: 0-521-80841-3
- [10] Collet, P.; Eckmann, J.-P. On the abundance of aperiodic behaviour for maps on the interval. *Comm. Math. Phys.* 73 (1980), no. 2, 115–160.
- [11] Dinaburg, E. I.; Sinai, Ja. G.: The one-dimensional Schrödinger equation with quasi-periodic potential. *Funkcional. Anal. i Prilozhen.* 9 (1975), no. 4, 8–21.
- [12] Fröhlich, J.; Spencer, T.; Wittwer, P.: Localization for a class of one-dimensional quasi-periodic Schrödinger operators. *Comm. Math. Phys.* 132 (1990), no. 1, 5–25.

- [13] Herman, M. R.: Construction d'un difféomorphisme minimal d'entropie topologique non nulle. *Ergodic Theory Dynamical Systems* 1 (1981), no. 1, 65–76.
- [14] Herman, M. R.: Une méthode pour minorer les exposants de Lyapounov et quelques exemples montrant le caractère local d'un théorème d'Arnold et de Moser sur le tore de dimension 2. *Comment. Math. Helv.* 58 (1983), no. 3, 453–502.
- [15] Johnson, R. A.: Ergodic theory and linear differential equations. *J. Differential Equations* 28 (1978), no. 1, 23–34.
- [16] Jorba, A; Tatjer, J. C.; Nunez, C.; Obaya, R.: Old and new results on strange nonchaotic attractors. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* 17 (2007), no. 11, 3895–3928.
- [17] Jakobson, M. V. Absolutely continuous invariant measures for one-parameter families of one-dimensional maps. *Comm. Math. Phys.* 81 (1981), no. 1, 39–88.
- [18] Jäger, T.: Strange non-chaotic attractors in quasiperiodically forced circle maps. *Comm. Math. Phys.* 289 (2009), no. 1, 253–289.
- [19] Jäger, T. : Strange non-chaotic attractors in quasi-periodically forced circle maps: Diophantine forcing. *Ergodic Theory Dynam. Systems* 33 (2013), no. 5, 1477–1501.
- [20] Kim, J.-W.; Kim, S.-Y.; Hunt, B.; Ott, E.: Fractal properties of robust strange nonchaotic attractors in maps of two or more dimensions. *Phys. Rev. E* (3) 67 (2003), no. 3, 036211, 8 pp.
- [21] Oseledec, V. I.: A multiplicative ergodic theorem. Characteristic Ljapunov exponents of dynamical systems. *Trudy Moskov. Mat. Obsc.* 19 (1968) 179–210.
- [22] Puig, J.; Cantor spectrum for the almost Mathieu operator. *Comm. Math. Phys.* 244 (2004), no. 2, 297–309.
- [23] Sinai, Y. G.: Anderson localization for one-dimensional difference Schrödinger operator with quasiperiodic potential. *J. Statist. Phys.* 46 (1987), no. 5-6, 861–909.
- [24] Sorets, E.; Spencer, T.: Positive Lyapunov exponents for Schrödinger operators with quasi-periodic potentials. *Comm. Math. Phys.* 142 (1991), no. 3, 543–566.
- [25] Wang, Y.; Zhang, Z.: Uniform positivity and continuity of Lyapunov exponents for a class of C^2 quasiperiodic Schrödinger cocycles. *J. Funct. Anal.* 268 (2015), no. 9, 2525–2585.
- [26] Wang, Y.; Zhang, Z.: Cantor spectrum for a class of C^2 quasiperiodic Schrödinger operators. *Int. Math. Res. Notices.* (to appear).
- [27] Young, L.-S.: Some open sets of nonuniformly hyperbolic cocycles. *Ergodic Theory Dynam. Systems* 13 (1993), no. 2, 409–415.
- [28] Young, L.-S.: Lyapunov exponents for some quasi-periodic cocycles. *Ergodic Theory Dynam. Systems* 17 (1997), no. 2, 483–504.

E-mail address: `bjerklov@kth.se`

DEPARTMENT OF MATHEMATICS, KTH ROYAL INSTITUTE OF TECHNOLOGY, 100 44 STOCKHOLM, SWEDEN